

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Validation of the 'Victim Empathy Response Assessment (ii)' (VERA-ii) with a Non-offending Adult Male Community Sample

Lister, Helen

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

This electronic theses or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Title: Validation of the 'Victim Empathy Response Assessment (ii)' (VERA-ii) with a Non-offending Adult Male Community Sample

Author: Helen Lister

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENSE AGREEMENT



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. <http://creativecommons.org/licenses/by-nc-nd/3.0/>

You are free to:

- Share: to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

VOLUME 1

Main Thesis and Service Evaluation Project

Helen Lister

Institute of Psychiatry, King's College London

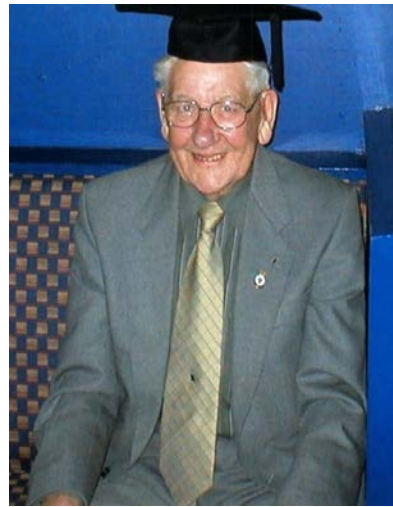
May 2012

A thesis submitted for the partial fulfilment of the degree of
Doctorate in Clinical Psychology

DEDICATED TO



Dad



Grandad

You shared the excitement at the start. If only you could share the celebration at the end.

This is for you x

ACKNOWLEDGEMENTS

I would like to thank and give credit to the following people without whom this journey would not have been possible.

Firstly, I would like to thank my research supervisor Professor Gisli Gudjonsson, for his support, advice and expertise, and ability to instil a calming influence upon me throughout the main research and write up process. I would like thank Dr Suzy Young for introducing me to the VERA(ii) and guiding me through the initial stages to get the project off the ground. Many thanks go to Gareth Hopkin and Emily Goodwin who were instrumental in the production of the VERA(ii), and the many tasks that they undertook to establish the project at Broadmoor. Also, to all at Third Dimension who produced VERA(ii). Additionally, a big thank you to Kelly Southgate for her unrelenting dedication to the recruitment of participants, and conducting the project, at Broadmoor despite the many obstacles encountered! Furthermore, for giving up her valuable time to provide inter-rater reliability. I would also like to thank Dr Nikki Collins for providing polygraph training and Professor Derek Perkins for sharing his expertise to assist with the analysis of the physiological data, and for his rescue efforts when locked in Broadmoor!

I am indebted to my clinical supervisors without whom I would not have gained the clinical skills and self confidence necessary to complete my training. Dr Rachel Mycroft, Professor Robin Morris, Dr Rachel Zwi, Dr Suzie Gratton, Dr Sarah Addison, Professor Myra Hunter and Dr Suzanne Haddad, I am extremely grateful to you all. Thank you all for sharing your knowledge and expertise and for your encouragement and guidance throughout. Additionally, I am immensely grateful to all the clients I have been privileged to work with, it would not have been possible without you. Thank you also to Dr Daniel Stahl for his statistical advice during the main research process.

I would like to thank all of the Academic and Administrative staff of the DClinPsy course, who thorough their friendly and supportive nature have created a secure base from which I have grown over the past three years. Special thanks to Dr Sue Rutter, Dr Laura Lunt and Carole Barnham for their warmth and understanding during difficult times.

A big thank you to all of course-mates – my 'IoP Family', whom I feel privileged to have shared the last three years, in-particular my housemate who made the increasingly frequent study weekends appear 'normal'. Thank you all for your friendship, kindness and the many many laughs....let's hope they continue!

Finally, thank you to my family and close others for their constant love, patience, encouragement and support, and for always believing in me. I would never have got this far if it wasn't for you all.

SUMMARY OF CONTENTS

MAIN STUDY:

Validation of the 'Victim Empathy Response Assessment (ii)' (VERA-ii)
with a Non-offending Adult Male Community Sample.....5

Supervised by Professor Gisli Gudjonsson

SERVICE EVALUATION PROJECT:

Use of the Mental Health Clustering Tool in Croydon Perinatal CMHT.....86

Supervised by Dr Rachel Mycroft

MAIN THESIS

Validation of the 'Victim Empathy Response Assessment (ii)' (VERA-ii) with a Non-offending Adult Male Community Sample

Helen Lister

Institute of Psychiatry, King's College London

Supervised by: Professor Gisli Gudjonsson

CONTENTS

ABSTRACT	10
1. INTRODUCTION	11
1.1. Empathy Deficits in Offenders	12
1.2. What is Empathy?	14
1.3. Where do Empathy Deficits Originate?	15
1.4. Measuring Empathy with Offenders	16
1.4.1. Global and Victim Empathy Self-Report Questionnaire Measures	16
1.4.1.1. Conceptualisation of Empathy	16
1.4.1.2. Limitations of Self-Report Questionnaire Measures	18
1.5. Development of Victim Empathy Measures	20
1.5.1. Victim Empathy Response Assessment (VERA)	21
1.6. Physiological Measures	23
1.6.1. Used With Sex Offenders	23
1.6.2. Physiological Measures of Emotion	24
1.7. Combining Physiological and Self-Report Measures of Emotion	27
2. RATIONALE	32
3. AIMS AND HYPOTHESES	34
3.1. Hypotheses	34
4. METHOD	35
4.1. Design	35
4.2. Power Analysis	35
4.3. Participants	35
4.4. Inclusion / Exclusion Criteria	35
4.5. Recruitment	35
4.6. Apparatus and Software	36
4.7. Materials	36
4.7.1. Victim Empathy Response Assessment (ii); (VERA-ii)	36
4.7.2. Wechsler Abbreviated Scale of Intelligence (WASI)	37
4.7.3. Eysenck Personality Questionnaire – Revised; Short Scale (EPQ-R)	37
4.7.4. Eysenck Impulsiveness, Venturesomeness & Empathy Questionnaire (IVE)	37
4.7.5. Maudsley Violence Questionnaire (MVQ)	37
4.7.6. Story Comprehension Task (SCT)	38
4.8. Procedure	38
4.9. Statistical Analysis	40

4.9.1. Physiological Data	40
4.9.1.1. Inter-rater Reliability	41
4.9.2. Data Treatment	41
5. RESULTS	43
5.1. Descriptive Statistics	43
5.1.1. VERA(ii)	43
5.1.2. Self-Report Measures	44
5.1.3. Physiological Measures	44
5.2. Hypotheses	45
5.2.1. Hypotheses 1(a) and 1(b)	45
5.2.2. Hypotheses 2 and 4	46
5.2.3. Hypothesis 3	46
6. DISCUSSION	48
6.1. Summary of Results	48
6.1.1. Hypothesis 1	48
6.1.2. Hypothesis 2	48
6.1.3. Hypothesis 3	49
6.1.4. Hypothesis 4	49
6.2. Interpretation of Results	50
6.3. Limitations, Theoretical and Clinical Implications	54
6.4. Strengths	57
6.5. Future Research	58
6.6. Conclusions	60
7. REFERENCES	62
8. APPENDICES	72

LIST OF TABLES

Table 5.1. Mean, SD, Medium and Range of Scores for Total Cognitive and Affective Empathy.....	43
Table 5.2. Wilcoxon test results for the difference between Total Cognitive and Affective Empathy.....	43
Table 5.3. Spearman's Rho Correlation of Total Cognitive and Affective Empathy.....	43
Table 5.4. Mean, SD, Medium and Range of Scores for Self-Report Measures, and IQ.....	44
Table 5.5. Mean, SD, Medium and range of Physiological Measures.....	44
Table 5.6. Wilcoxon test results for the difference between Mean Baseline and Total Mean for both SRL and HR.....	45
Table 5.7. Reliability (Cronbach's α) for VERA(ii) Cognitive and Affective Empathy Subscales for each vignette and total vignettes.....	45
Table 5.8. Item Analysis for VERA(ii) Cognitive and Affective Empathy Subscales.....	45
Table 5.9. Spearman's Rho Correlations of Cognitive and Affective Empathy with Self-Report Measures, and IQ.....	46
Table 5.10. Spearman's Rho Correlations of Cognitive and Affective Empathy with Physiological Measures.....	46
Table 5.11. Spearman's Rho Correlations between Mean Total SRL and Mean Total HR.....	47
Table 5.12. Spearman's Rho Correlations between Total mean Change SRL and Total Mean Change HR.....	47

APPENDICES

Appendix 1.	Participant Information Sheet.....	73
Appendix 2.	Ethical Approval Letter.....	75
Appendix 3.	Ethical Approval Letter – Substantial Amendment.....	80
Appendix 4.	VERA(ii) Affective Empathy Questionnaire.....	83
Appendix 5.	VERA(ii) Cognitive Empathy Questionnaire.....	84
Appendix 6.	Example of Physiological Recording Output.....	85

ABSTRACT

Developing empathy has become a central component of cognitive behavioural treatment programmes for offenders, who it is argued have victim empathy deficits. It is therefore important to evaluate the effectiveness of these interventions. Research is moving towards utilising victim empathy measures as the most effective tools for this purpose, although evidence has been mixed regarding their psychometric properties. The Victim Empathy Response Assessment (VERA) was developed as a victim specific measure to assess cognitive and affective empathy in forensic settings. Limited validity has been demonstrated using an adult male forensic and community sample. Following participant feedback suggesting how to increase the validity of the tool a second video prototype VERA(ii) was developed. This study aimed to evaluate the validity of the VERA(ii) with 51 non-offending adult community males. Participants viewed the VERA(ii) whilst simultaneously providing physiological measurements of skin resistance level (SRL) and heart rate (HR). Following this they completed the VERA(ii) cognitive and affective empathy questionnaires, also the Wechsler Abbreviated Scale of Intelligence, Eysenck Personality Questionnaire Revised - Short Form, Eysenck Impulsiveness, Venturesomeness and Empathy Questionnaire, Maudsley Violence Questionnaire and the Story Comprehension Task. Wilcoxon signed-ranks tests demonstrated a significant difference between cognitive and affective empathy and SRL from mean baseline to mean total SRL after viewing the VERA(ii). However, Spearman's Rho Correlations found no significant associations between physiological measures and total cognitive or affective empathy. A significant negative correlation was found between cognitive empathy and MVQ Machismo, and a significant positive correlation was found with the Story Comprehension Task. No other significant associations were demonstrated. Overall, the results of the current study have provided limited evidence for the validity of the VERA(ii). Limitations of the study have been highlighted and discussed with regards to theoretical and clinical implications. Recommendations for future research have subsequently been presented.

The original intention of this research was to validate the VERA(ii) using a non-offending adult male community sample and adult male in-patient offenders. However, unfortunately due to practical difficulties severely limiting the recruitment of the offending group this part of the study could not be fully completed. The study therefore focuses upon a community sample only.

1. INTRODUCTION

Baron-Cohen describes empathy as; “one of the most valuable resources in our world” (Baron-Cohen, 2011, p. 124), a lack of which he asserts can explain human cruelty at various degrees for example, ranging from upsetting another by saying something insensitive to the extremes of committing rape or murder. Without empathy he argues, the risk of hurting others, destroying relationships and causing conflict, both at individual and societal levels, is great. It is therefore essential that we attempt to gain an understanding of empathy to be able to address these risks.

In his book ‘Zero Degrees of Empathy: A New Theory of Human Cruelty’ he presents his theory providing scientific evidence to support his ideas. Laying the foundations by describing empathy as, “our ability to identify what someone else is thinking or feeling, and to respond to their thoughts and feelings with an appropriate emotion” (Baron-Cohen, 2011, p. 11), he then builds his theory. To begin with, he describes empathy as falling upon a normally distributed continuum ranging from zero to six degrees. At zero degrees one is seen as being completely devoid of empathy for example, psychopaths (zero degrees negative) or individuals with autism (zero degrees positive), whereas at six degrees people are deemed highly empathic. Presenting evidence from brain scans he argues that the brain contains an ‘empathy circuit’ present in all individuals. However, it is from abnormalities in this circuit that individual differences arise in the degree of empathy one has. Abnormalities in the circuit may be present as a consequence of genes, early care-giving experiences or current stresses, and it is for these reasons that empathy deficits can be either permanent or temporary. For example in a ‘normal individual’ the empathy circuit may temporarily shut down at the very moment of engaging in an unempathic act e.g. verbal abuse, to be restored after this temporary period. Whereas for individuals with personality disorders such as someone who is ‘Type B – Borderline’, ‘Type P – Psychopathic’ or ‘Type N – Narcissistic’ the empathy circuit maybe permanently switched off.

Focusing upon zero degrees of empathy, specifically zero degrees negative, Baron-Cohen argues that uncaring, and at times horrendous acts of human cruelty, committed by individuals towards others should not deem that individual to be innately ‘evil’, unable to be changed following interventions, but instead should be seen as an individual suffering from a deficit in their empathy circuit, be it permanent or temporary. He argues that if we re-conceptualise ‘evil’ in this way, this opens up the possibility for understanding such individuals and being able to develop interventions to assist in overcoming their deficit. In summary, he argues it is

imperative to continue to focus upon empathy, integrating the ideas put forward in his theory into research and treatment at all levels upon the continuum.

Taking this on board, this study is concerned with the validation of a new measure of victim empathy – the Victim Empathy Response Assessment (ii) (VERA(ii); developed by Young et al.) The VERA(ii) was developed to provide an ecologically valid experimental measure of ‘cognitive’ and ‘affective’ empathy in relation to specific categories of criminal offence, for the purpose of identifying treatment needs of offenders and evaluating treatment outcome. The VERA(ii) consists of five videos in which victims, played by actresses, are interviewed about their experience of different incidents. Participants are requested to watch each interview after which they are asked to rate how they felt while watching the interview (affective empathy) and how they believe the victim might feel about their experience (cognitive empathy). VERA(ii) was produced as an updated version of the original VERA which consisted of audio ‘radio interviews’ (Young, Gudjonsson, Terry, & Bramham, 2008). Validation of VERA has shown good internal reliability of the affective and cognitive constructs and it has shown partial construct validity against existing empathy self-report questionnaires (Terry et al., 2009). However, participants suggested that the audio format of the VERA seemed artificial and it was suggested that this may be improved if a visual video format was introduced; VERA(ii) therefore was produced and needs to be validated. As a means to address some of the limitations of current empathy measures this study aims to employ physiological measurements alongside self-report measures as an empirical measurement of empathy. It specifically aims to measure the emotional distress of participants whilst watching the videos as research has shown that this is an indication of the empathic process occurring.

1.1. Empathy Deficits in Offenders

Empathy has been described as, “The ability to perceive another person’s point of view, the ability to experience the perspective and role of others, and the capacity to behave compassionately (i.e. ‘affective responsiveness’)” (Fisher & Howells, 1993 as cited in Young, Gudjonsson, Terry, & Bramham, 2008, p.191-2). It plays an important role in social relationships as it helps people to understand, anticipate and respond appropriately to the feelings and behaviours of others in a variety of contexts. Consequently having empathy deficits maybe socially disabling to the individual (Baron-Cohen, 2011). The ability of an individual to empathise is therefore positively correlated with pro-social or altruistic behaviour (Batson, Fultz, & Schoenrade, 1987). For this reason empathy development is a commonly defined treatment goal for sexual offenders (Freeman-Longo, Bird, Stevenson, & Fiske, 1995), who are generally regarded as having deficits in victim empathy (Pithers, 1994; Varker, Devilly, Ward, & Beech, 2008), with many experts recommending empathy training as part of any programme that attempts to reduce recidivism among offenders (Regher & Glancy, 2001).

The rationale driving empathy research with this population is that offenders, including; rapists, child molesters and abusers (Goleman, 1995; Wiehe, 1997; Fisher, Beech, & Brown, 1999)

have deficits in empathy which results in them behaving in undesirable ways (Fernandez & Marshall, 2003; Marshall, Hamilton, & Fernandez, 2001) (i.e. facilitating offending behaviour by engaging in anti-social behaviour, particularly violent and sexual offences; Farrington, 1998). For example, when applying this to sexual offenders Marshall and Barbaree (1990) explain that the low levels of empathy in this population contributes to their offending by disinhibiting their sexual arousal. This is accounted for theoretically when they argue that the sexual offender can commit his offence because his low empathy reduces either his understanding of the physical and emotional suffering of his victim, or instead because he simply does not care (Varker et al., 2008), in contrast to a non-offender with no empathy deficits. This explanation, that deficits in one's empathy inhibits one's understanding of another's experience, thus how one interacts with another, has also been applied to psychopaths who Harpur, Hakstian, and Hare (1988) report are more violent towards others than non-psychopaths. This explanation has also been applied to individuals with 'theory of mind deficits' such as individuals on the autistic spectrum (Baron-Cohen, 2011) and those who have schizophrenia (Langdon, Coltheart, & Ward, 2006).

Ryan (1999) states that empathy provides the highest deterrence for abusive behaviour towards others. Ward, Keenan, and Hudson (2000) reported the ability of one to empathise is influenced by heightened emotional states including anger, profound loneliness and sexual arousal. Consequently, Mulloy, Smiley, and Mawson (1999, p.16) state that, "some form of empathy training is, therefore, a common treatment component of those convicted of crimes such as assault, robbery, murder, and sexual assault." In the sex offending population developing empathy has become a central component of cognitive behavioural treatment programmes, including relapse prevention (Marshall, 1999). Knopp, Freeman-Longo, and Stevenson (1992) present statistics from a US National Survey (cited by Beech, Fisher, and Beckett, 1998) that 94% of sex offender programmes in North America include victim empathy as a treatment goal, with Freeman-Longo et al. confirming in 1995 that a component of empathy training with adult sexual offenders is indeed being delivered within these treatment programmes. They explain that the inclusion of empathy as a treatment component originated from the belief that if offenders can develop empathy for their victims this will then have an inhibitory effect upon their motivation to offend. There are also, as presented by Varker et al. (2008), important economic benefits to be derived from appropriate and effective treatment programmes. Furthermore, Pithers (1994) argues, it will improve the offenders capacity for intimacy in interpersonal relationships and contribute to maintaining their motivation to change. An argument supported by Briggs, Doyle, Gooch, and Kennington (1998) who state that the presence of empathy in an offender allows the offender to understand the impact of his/her actions upon their victims, subsequently allowing the offender to feel remorse. With this in mind victim empathy work within cognitive behavioural programmes have incorporated narratives from the victims point of view as a means to increase the empathic responding of the offender by developing an understanding of the emotions experienced by victims.

1.2. What is Empathy?

Empathy has commonly been described as a multi-component process (Marshall, Hudson, Jones, & Fernandez, 1995; Marshall, 2011). Following a review of the literature on empathy, including a specific emphasis towards research with sex offenders, Marshall and Marshall (1995) developed a process model of empathy, stating that empathy consists of four components or stages which must be achieved consecutively for the empathic process to be complete. The four stages include: (1) Emotion Recognition, (2) Perspective-taking, (3) Emotion Replication and (4) Response Decision. They explain that for emotion recognition to occur one must be able to accurately recognise the emotional stage of an observed other. Perspective-taking is achieved when one places oneself in the observed person's place; to be able to take the other's perspective. Once this has been achieved emotion replication occurs whereby one has an emotional response matching, or almost matching, that to which the observed other is experiencing. Finally, one makes a response decision; deciding how one is to react to the observed other in response to the feelings that have been generated.

Marshall and Marshall (2011) following a review of the concept of empathy and the evolution of models used to account for the empathic process report that current research, including the development of tools to measure empathy, are based upon the understanding that empathy contains both cognitive and affective components. To explain; 'cognitive empathy' is the ability to understand another person's feelings, being able to imagine their perspective and predict how they may feel as a result. 'Affective empathy' on the other hand is experiencing an appropriate change in mood in response to another's; ultimately sharing the emotional experience of another, for example feeling distressed by another's unhappiness. They differentiate this from sympathy when they state that sympathy involves feelings of concern. However they state that empathy and sympathy can co-occur, also empathy can precede feelings of sympathy (Feshback, 1975; Gladstein, 1983; Moore, 1990; Marshall et al., 1995; Kerem, Fishman, & Jorrelson, 2001). Baron-Cohen (2011) when describing his theory of six degrees of empathy gives the following examples of individuals in relation to cognitive and affective empathy. For example, as one moves up from zero degrees of empathy an individual will display both cognitive and affective components. However, at zero degrees negative where the psychopath is situated, evidence has shown only the cognitive but not the affective component to be intact. Finally, for the individual with classic (low functioning) autism situated at zero degrees positive, both cognitive and affective elements are typically lacking. Together cognitive and affective empathy would therefore correspond to the first three stages of Marshall and Marshall's process model.

Empirical research with offenders and non-offenders has led to the suggestion that empathy can be further divided into two types; namely 'global empathy' and 'victim empathy'. Where empathy deficits exist, global empathy suggests that offenders have a generalised empathy deficit in comparison to non-offenders, such they are unable to empathise over a range of

contexts and/or to persons; Baron-Cohen (2011) highlights the 'Psychopath - Type P' as an example of this. Victim empathy in comparison suggests that offenders might have particular empathy deficits in relation to specific offending situations including; types of offences and types of victims, rather than possessing general empathy deficits; an argument presented by Marshall et al. in their 1995 paper 'Empathy in Sex Offenders'. Overall, despite most research focusing upon global empathy, few empirical studies have supported the assumption of a generalised empathy deficit in offenders in comparison to non-offenders (Marshall, Anderson, & Fernandez, 1999).

Evidence for victim empathy has been reported by Farrington (1998), among others (Curwen, 2003; Tierney & McCabe, 2001), who presented findings showing an association between a deficit in empathy and violent and sexual offences, specifically among paedophilic sex offenders. McGrath, Cann, and Konopasky (1998) also suggest that sex offenders are not particularly deficient in general empathy but are deficient in their empathy for non-specific victims of child sexual abuse, such that they are deficient in recognising the emotions of this group and those of their own victims but not the emotions of others generally. In addition there is also evidence to suggest that offenders may have deficits in cognitive empathy but not affective empathy (Fisher & Howells, 1993; Jolliffe & Farrington, 2004). In summary, the evidence suggests that the empathy/offending relationship is complex; that empathy is not a stable trait but fluctuates dependent upon the context; type and / or location of the offence, the level of distress created and the person (i.e. the victim and / or the offender). Young et al. (2008) have therefore argued that this must be considered in the effective measurement of empathy among offenders.

1.3. Where do Empathy Deficits Originate?

Baron-Cohen (2011) has argued that empathy deficits may be permanent or temporary and are caused by abnormalities in the empathy circuit in the brain. Marshall and Marshall (2011), based on their process model of empathy, have made suggestions about where they believe empathy deficits may originate, addressing each stage of the model in turn. For example, a deficit in empathy would emerge at stage one: emotion recognition if the observer is unable to recognise the emotions displayed by another, which they argue will prevent the empathic process from ever beginning. Should one be able to accurately recognise the emotions of the observer but not be able to understand why the other is experiencing such emotions, therefore limiting one's ability to take on the observer's perspective at stage two, possibly as a result of one having poor theory of mind skills, the empathic process would again be impaired. Alternatively they argue if one recognises the distress in the observer but has an incongruent response (i.e. one is pleased by the others distress as a sadist would, then one would feel no urge to respond appropriately to ease the other's distress). Finally, if one successfully achieves emotional recognition of the other and is able to take the other's perspective yet by doing so this causes the individual to become so overwhelmed by the others distress that the individual

needs to self-soothe, then one will be unable to complete the empathic process by failing to respond to, and ease, the other's distress.

1.4. Measuring Empathy with Offenders

As Marshall and Marshall (2011) argue if the empathic process has been achieved one will have experienced a resulting emotional response. Rosenthal, Gratz, Kosson, Cheavens, Lejuez, and Lynch (2008) state that emotional responses are made up of several distinct components including; subjective experience, physiological arousal and overt motoric behaviour. It is therefore possible to measure empathy using these components. Techniques used to achieve this have included; applying clinical ratings, self-report measures (Jolliffe & Farrington), measuring facial responses (Marcus, Roke, & Bruner, 1985), and measuring physiological reactions (Eisenberg & Fabes, 1990) as well as functional magnetic resonance imaging (fMRI; de Vignemont & Singer, 2006). It has been suggested that self-report measures provide a useful method for assessing the subjective component of emotional responses whereas measures of physiological arousal allow for a more complex and thorough objective understanding of emotional responding (Rosenthal et al., 2008). Despite this measuring empathy with offenders has typically relied on self-report measures. Jolliffe and Farrington (2004) in their review of empathy and offending confirmed this reporting that the majority of studies assessing empathy among offenders employ self-report questionnaires, with too few studies adopting alternative methods. A criticism also highlighted by Grady, Broderson, and Abramson (2011) when evaluating the psychological measures available for use with sex offenders, be they general or victim specific.

As mentioned self-report measures are useful tools for assessing the subjective component of emotional responses, including empathy. They are easy to administer and have the advantage of being relatively quick to complete. However, they also have a number of limitations, particularly when applied to an offending population. A discussion of which now follows.

1.4.1. Global and Victim Empathy Self-Report Questionnaire Measures:

1.4.1.1. Conceptualisation of Empathy

Historically the assessment of empathy has relied mainly on global empathy measures, those measuring a person's general level of empathy. Given the lack of evidence of a generalised empathy deficit in offenders, in comparison to non-offenders, it has been argued that it is perhaps not surprising that studies measuring general empathy have been unreliable in their ability to distinguish between groups for example, in their ability to distinguish sex offenders from other populations (Grady et al., 2011). This has resulted, more recently, in the development of victim specific empathy measures that have been designed for use with specific types of offenders.

Grady et al. (2011) in their article, 'The state of psychological measures for adult sexual offenders' reviewed the instruments commonly available to measure the core treatment targets in sexual offender treatment programmes, including empathy. When reviewing measures of victim empathy they reported numerous limitations, placing particular emphasis on the finding that although there are a range of instruments available the results of studies assessing the psychometric properties of these instruments (i.e. those with proven reliability and validity) are inconsistent. They argue that this is mainly a consequence of the instruments measuring different aspects of the empathy construct which, Serran (2002) suggests is explained partly due to the lack of agreement among researchers to define empathy; supported by Young et al. (2008) and Jolliffe and Farrington (2004), a common criticism pertaining to global empathy measures also.

The lack of a consistent definition of empathy has led to discrepancies within research findings, occurring with both global and victim empathy measures. For example, some empathy measures are based on definitions of cognitive empathy only, such as the Child Empathy Test (CE; Hanson & Scott, 1995), whereas others are based solely on definitions of affective empathy such as the Impulsiveness, Venturesomeness and Empathy measure (IVE; Eysenck & Eysenck, 1978), yet others are based on definitions containing both cognitive and affective empathy, such as The Child Molester Empathy Measure (CMEM; Fernandez, Marshall, Lightbody, & O'Sullivan, 1999). This has repercussions for the findings of studies that have employed either general empathy measures or those measuring either cognitive or affective empathy in isolation. Consequently this gives rise to mixed results, for example as mentioned above, Fisher, Beech, and Browne (1999) have presented findings suggesting that offenders may be deficient in cognitive empathy but not affective empathy. However, Scully (1988) presented conflicting evidence when assessing self confessed rapists. Here it was demonstrated that over half of the group displayed cognitive empathy towards a female victim (i.e. they were able to understand the victims feelings and correctly classify her reactions) yet over half of this group, when asked about their feelings towards the victim (affective empathy), were unable to identify with these instead indicating that they felt nothing towards the victim.

A systematic review and meta-analysis of measures relating to cognitive and affective empathy in relation to offending, incorporating both global and victim empathy measures, concluded that low cognitive empathy was strongly related to offending with low affective empathy being weakly related (Jolliffe & Farrington, 2004). When reviewing the data assessing for cognitive and affective empathy deficits in offenders compared to non-offenders they also reported that the relationship between low empathy and offending was relatively strong for violent offenders yet relatively weak for sex offenders. However, the relationship between low empathy and offending was not significant after controlling for other variables namely, intellectual functioning and socio-economic status (SES).

1.4.1.2. Limitations of Self-Report Questionnaire Measures

Tierney and McCabe (2001) identify three broad problems with self-report questionnaires. Firstly, they argue that the psychometric properties of self-report questionnaire measures designed to measure empathy in offenders, among other constructs, vary tremendously. They qualify this by presenting a range of evidence demonstrating that some measures have no established validity or reliability yet others, although presenting evidence of both validity and reliability, nevertheless could improve this by using more specific validation methods and larger sample sizes. For example, when reviewing the Empat A (McGrath et al., 1998) they report that the measure has good internal reliability but highlight that its test-retest reliability was not assessed and it was unable to discriminate between groups thus raising doubts about the discriminant validity of the measure. Conversely, when presenting evidence for the Victim Empathy Distortions Scale (VEDS; Beckett & Fisher, 1994) they conclude that the measure demonstrates good levels of internal reliability, test-retest reliability and discriminant validity. However, despite this they argue that the measure requires further evaluation before it can successfully prove that it measures what it claims to do so i.e. the offenders empathy for his victims. In order to do this they suggest that the measure is evaluated against existing measures which profess to measure the same construct, although the difficulty with this they discuss is the lack of such tools that can act as benchmarks, as has been highlighted previously.

The second problem identified by Tierney and McCabe (2001) of self-report questionnaire measures used with offenders is identifying which is the best measure to use. As mentioned previously, inconsistencies with regards to the definition of the concept (i.e. is the measure a general or victim empathy measure and is it measuring the cognitive or affective aspects of empathy, or both?). Additionally, was the measure assessed using an offender or non-offending population? To date studies have produced mixed, conflicting findings, for example among research with sex offenders (Pithers, 1994; McGrath et al., 1998). Coupled with the lack of psychometric evaluation of the measures available this has the outcome of making the data generated from them questionable therefore begging the question what are we truly learning about offenders from these measures?

The third problem of self-report measures they argue is the general 'transparency' of the purpose of the measure (i.e. it is often apparent what construct is to be measured which, it has been argued, makes them susceptible to response bias or 'faking good' by offenders who, rather than answering truthfully, instead give socially desirable responses). Hanson and Scott (1995) argue that this is because even if sexual offenders fully admit to their offences they nevertheless bow to the strong social pressure for them to generate acceptable explanations for their offences. As such Anastassiou-Hadjicharalambous and Warden (2007) conclude that self-report information may instead inform the clinician more about how the individual wants to see oneself, or to be seen by others, than how one actually feels during an empathy-inducing context. Alternatively an offender may strive to give socially desirable responses in the hope

that by doing so he will be deemed to have been successfully rehabilitated, reducing his level of risk which may see him moved to a lower security setting, for example.

When 'faking good' occurs it is impossible to accurately determine if the answers given by the respondent reflect their actual opinions or an attempt to be perceived in a socially desirable manner. (Furnham, 1986; McGrath et al., 1998; Tierney & McCabe, 2001; Young et al., 2008; Grady et al., 2011). This has numerous consequences specifically with regards to the offender population in terms of assessment of risk, risk management, treatment planning and designing and analysing the effectiveness of treatment programmes. Furnham (1986) explains that in an attempt to detect respondents who choose socially desirable responses, some measures have incorporated lie-scales as demonstrated by Eysenck and Eysenck (1975) in the Eysenck Personality Questionnaire. Alternatively some researchers have employed qualitative approaches as an attempt to overcome this, for example using offenders' life stories to examine underlying schema. However, this method also has its limitations (Grady et al., 2011).

An alternative view to the 'faking good' hypothesis applied to offenders when completing self-report questionnaire measures is presented by Eisenberg and Fabes (1990). Instead of an offender consciously 'faking' his response to appear socially desirable to others they suggest that in order for one to be able to accurately complete self-report measures one must first be aware of what one is feeling and be able and willing to report these feelings accurately. To expand, they pose the question 'is one able to decipher and accurately communicate one's emotional states?' (Eisenberg & Fabes, 1990, p. 135). Supported by Marshall and Marshall (2011) in relation to their process model of empathy as discussed previously, if one is unable to recognise emotion in oneself or in others such that one cannot complete the empathy process, how then can one report on their feelings of empathy? Furthermore Baron-Cohen (2011) argues that relying on self-report measures of empathy can be problematic as an individual may simply believe they are more empathic than they actually are. He supports this by explaining that a person with poor empathy is often the last person to realise that this is so.

Given the critiques of global versus victim empathy measures, the conceptualisation of the empathy construct and the limitations of the dominant self-report questionnaire based method of assessment with offenders, it has been concluded that "better measures of empathy are needed" (Jolliffe & Farrington, 2004, p. 441). Marshall et al. (1995) suggested one way to address this when researching empathy with offenders is for researchers to abandon global measures of empathy in favour of victim empathy measures. Supported by Tierney and McCabe (2001) who, after evaluating the utility of self-report measures among paedophilic sex offenders, concluded that measures of victim empathy are of greater clinical utility than measures of general empathy. Consequently there has been a shift away from measures of general empathy to focus instead on those that are 'victim specific' that measure empathy deficits specific to particular victims or situations, which by default also makes these specific to offender type (Fernandez & Marshall, 2003).

1.5. Development of Victim Empathy Measures

With the development of victim empathy measures most researchers have attempted to move away from the dominant self-report questionnaire format. This is because as the measures are becoming more specific to types of offenders and their victims more detailed information than can be given in a simple questionnaire format, is required to help to elicit attitudes / emotional responses that form the basis of empathy, from offenders. Research has therefore adopted the use of written vignettes to be used in combination with self-report questionnaires. West (1982) suggested vignettes to be a superior method of eliciting attitudes compared to questionnaires alone. Therefore they should provide a more detailed and accurate assessment of an offenders victim empathy if employed. For example, Hanson and Scott (1995) developed the Child Empathy Test (CE) and the Empathy for Women Test (EFW; Hanson & Scott, 1995) to assess cognitive empathy deficits in sexual offenders.

In the CE participants read 16 vignettes describing a range of interactions depicting instances of childhood sexual abuse and non-abusive interactions, after which they are asked to rate how the child felt on a Likert scale from 1-7 where 1 = 'very upset, unhappy, or scared' and 7 = 'very good, happy, cheerful'. Similarly in the EFW participants read 15 vignettes relating to situations associated with sexually abusive and non-abusive interactions between men and women, after which they are asked to identify which of the scenarios would be felt as sexually threatening by the woman, as well as rating the feelings of the woman. Unfortunately this measure was reported to have poor reliability (internal consistency: $\alpha = 0.30$) and discriminant validity (Hanson & Scott, 1995). However, the authors acknowledged that the measure focused mainly on cognitive empathy and suggested this may have been improved if additional affective empathy items were incorporated. Furthermore, they suggested that increasing the number of different feeling states for respondents to rate may improve its reliability although they stressed that this may leave the measure open to self-presentation biases. Finally, another suggestion to increase the reliability of the measure was to increase the number of vignettes, to counteract the increased length of time it would then take to complete the measure they suggested that the vignettes could be presented in an audio or visual format making it quicker to complete and also easier for the offenders to respond.

More recently victim specific measures assessing both cognitive and affective empathy have included the Child Molester Empathy Measure (CMEM; Fernandez, Marshall, Lightbody, & O'Sullivan, 1999), designed for use with sexual offenders of child victims and the Rapist Empathy Measure (REM; Fernandez and Marshall, 2003) designed for use with sexual offenders of adult victims. Both measures built upon the method employed by Hanson and Scott (1995) and consist of written vignettes describing a child or adult who has been the victim of a road accident (non-offending theme), a victim of sexual abuse over a period of time, and the offenders own victim (sexual offending theme). The participant is requested to read each vignette after which he is required to indicate on a Likert scale from 0-10 his recognition of the distress and harm caused to the victim and his own emotions towards the victim. Initial findings

of the CMEM demonstrated good validity; for sex offenders internal reliabilities of the subscales ranged from $\alpha = 0.80 - 0.94$ (including sexual abuse, car accident and own victim subscale); non-sexual offender comparison $\alpha = 0.84 - 0.85$ (sexual abuse / car accident subscales only). Test-retest reliability for non-sexual offender subscales found; $r = 0.64 - 0.83$. However, relatively small sample sizes were used. Good validity was also found for the REM when used with a student sample, with internal consistency of the accident victim scale and sexual abuse victim ranging from $\alpha = 0.85 - 0.93$ and $\alpha = 0.90 - 0.91$ respectively, from first to second testing. Test-retest reliability for the scales demonstrated $r = 0.81$ and $r = 0.84$ respectively. Unfortunately, poor discriminant validity was found when used with a sample of adult sexual offenders.

In summary, as suggested by Jolliffe and Farrington (2004) victim empathy measures are being created. However, further work is required in the validation and continued development of these measures with the offending population. Most recently, Young, Gudjonsson, Terry, and Bramham (2008) designed a victim empathy measure, attempting to address the limitations of existing victim empathy measures as described and incorporating the suggested improvements of previous studies.

1.5.1. Victim Empathy Response Assessment (VERA)

Young et al. (2008) developed the Victim Empathy Response Assessment (VERA) with the "aim of creating an ecologically valid measure of cognitive and affective empathy in relation to specific categories of criminal offence" (p. 193). Their rationale being; to create a reliable and valid tool of victim empathy to be used in both research and clinical assessment capacities within forensic services. VERA attempted to overcome limitations of existing victim empathy measures by not only using vignettes (in addition to questionnaires), which have been reported to be superior when attempting to elicit attitudes compared to questionnaire-only based measures (West, 1982), but to deliver these using a radio-interview format which is considered a more realistic assessment of victim empathy when compared to written vignettes (Hanson & Scott, 1995).

The VERA is therefore comprised of five staged tape-recorded vignettes presented as radio interviews between a female 'victim' and a male interviewer. The same male interviewer conducts all five interviews, with each interview consisting of a different female 'victim'. The themes of the interviews differ such that two are non-offending (a car accident and an accidental house fire), two are of a sexual offending nature (sexual assault and child sexual abuse) and one focuses on violence (assault). Participants listen to each interview after which they are requested to complete two questionnaires; the first requesting them to rate, using a five-point Likert scale (0 = 'Not at all'; 4 = 'Very much so'), how 'they' felt while listening to the interview (affective empathy rating) and the second asking them to rate how they believed 'the victim' felt (cognitive empathy rating) when talking about her experience. Each questionnaire consists of the same 13 emotions; nine negative (worried, sad, upset inside, distressed, sorry,

disturbed, angry, disgusted and frightened) and four positive (thrilled, interested, happy and excited).

The VERA was based upon the contemporary definition of empathy as being a two component construct encompassing 'affective empathy' (experiencing an appropriate change in mood in response to another's) and 'cognitive empathy' (the ability to understand another person's feelings/perspective taking) (Gladstein, 1983; Moore, 1990; Marshall et al., 1995; Kerem, Fishman, & Jorrelson, 2001). Also upon the view presented by Marshall et al. (1995), when conducting empathy research with offenders, that offenders may display specific empathy deficits (i.e. either with particular types of victims or particular types of offence rather than have deficits in general empathy). The VERA was therefore developed to incorporate these concepts thus a measure of both affective and cognitive empathy was produced to address different categories of offence; specifically sexual offences (sexual assault and child sexual abuse) and a violent offence (assault).

The VERA has been validated in two studies to date using an adult male forensic inpatient sample (Young et al., 2008) and an adult male community sample (Terry, Gudjonsson, & Young, 2009). Factor analysis of the scale using a forensic sample initially revealed two distinct factors; affective and cognitive empathy, with a high level of internal reliability across subscales (Cronbach's alpha ranging from 0.85 to 0.94), also replicated with a community sample (Cronbach's alpha ranging from 0.90 to 0.94). Further data analysis by Terry et al. (2009) provided little evidence of concurrent validity of the tool. However, this was only found in relation to the non-offending vignettes and was only significant in relation to affective empathy. Thus Terry et al. (2009) concluded that their findings were indicative of a relationship between general empathy measures and affective, rather than cognitive empathy, and to non-crime related scenarios. Finally, the data analysis was unable to demonstrate discriminant validity as the VERA scores were significantly lower for the community sample than the forensic sample and failed to confirm the predicted differences in levels of empathic response when comparing offending and non-offending scenarios. Further investigation into the validity of the tool is therefore required.

Future developments of the VERA, suggested by Young et al. (2008) following qualitative participant feedback from these studies, included attempting to enhance the 'realism' of the interviews by producing these in visual format as the audio format was reported to feel artificial. Also, using professional actors and liaising with victim support groups as to the design of the interview content were suggested, as a means to enhance the reliability of the measure. Reacting to these suggestions VERA(ii) was produced, updating the audio interviews to a television interview format which now needs to be validated.

Terry et al. (2009) when attempting to explain why VERA scores did not correlate with intellectual functioning and personality variables as hypothesised suggested this may be so if as suggested previously, general empathy and victim empathy are independent constructs.

Additionally, owing to the lack of alternative victim specific measures from which to validate new measures this begs the question, 'how then does one assess construct validity for victim empathy measures?' The American Psychological Association (APA; 1986, as cited by Tierney & McCabe, 2001) state that where test developers are assessing the validity of their measures they should do this using a variety of different validity assessment procedures. Taking this in to consideration, and in light of the unexpected findings with regards to assessing the validity of VERA by Terry et al. (2009), one may look to measure the objective components of empathy related emotions i.e. physiological measurements, alongside the self-reported subjective components (Rosenthal et al., 2008).

1.6. Physiological Measures:

1.6.1. Used With Sex Offenders

Research focusing upon deviant sexual arousal with sex offenders has shown physiological measurements to be superior to self-report measures, with sex offenders significantly underestimating levels of deviant arousal using the latter method (Laws, 1989; Murphy & Barbaree, 1988; Pithers & Laws, 1989; O'Connell, 2000). Typically there have been two ways of utilising physiological measures with sex offenders. Firstly, the traditional approach using the Penile Plethysmograph (i.e. using offenders' erections to measure their erotic preferences; Laws, 1989), and secondly, the use of the lie detection approach (i.e. polygraph testing; Grubin, 2002). It is highlighted that these approaches are fundamentally different, the former of which is used less nowadays. The latter, polygraph testing, is one way of attaining physiological measurements, a method which is used increasingly in the clinical setting in the treatment, management and supervision of sex offenders (Grubin & Madsen, 2006; British Psychological Society, 2004) as it allows important information to be elicited from the offender that is otherwise difficult to obtain (O'Connell, 2000). For example, polygraph testing has typically been used in research with offenders as a means to detect deception (i.e. when an individual is attempting to conceal the truth). The theoretical underpinning being simply that physiological activity increases when one is emotionally aroused / distressed in response to a stimulus. Comparing the rate of physiological activity at a particular point in time to the activity at another point in time allows one to detect any changes (i.e. the individual's level of arousal / distress in relation to particular stimuli).

Referring again to the research and treatment of sex offenders, polygraph testing has been argued to be an invaluable method for obtaining important information with regards to an offender's sexual history, specifically to encourage the offender to give an accurate and truthful account. As Grubin stated after completing a relatively small study assessing the utility of the polygraph with sex offenders in 2002; the most effective use of the polygraph may therefore be as a 'truth facilitator' rather than a 'lie detector' (Grubin, 2002, p. 51). Fundamentally polygraph testing is employed to encourage offenders to make clinically important disclosures that are relevant to their treatment and supervision rather than as a means to make definitive

judgements of truth. This has the advantage of bringing worrying behaviours to the attention of clinicians in time for effective intervention and additional treatment to be introduced before re-offending occurs. (Grubin, 2002). The rationale underpinning polygraph testing is based on the argument that the offender will disclose more truthful information as he fears that in not doing so his physiological responses being measured will inform others of his deceit (i.e. psychological states that the offender may otherwise be unwilling to report). As Blascovich, Vanman, Berry Mendes, and Dickerson (2011, p. 3) summarise when they say, “criminals often confess to a crime if they believe someone has insight into their ‘true’ thoughts as revealed via autonomic nervous system measures”.

Polygraph testing has been proven as a useful tool to assist with treatment planning and monitoring participation in treatment post-conviction as well as for its original use pre-conviction (Grubin & Madsen, 2006). As is used in lie-detection polygraph testing (measuring an individual’s physiological responses to stimuli) it can also be an effective tool to assess for the veracity of an individual’s responses when completing self-report measures (i.e. emotional / attitudinal). Blascovich et al. (2011) explain why this is so, arguing that self-reported attitudes represent only the part of the individual’s belief structure that one is willing to share with others. Physiological reactions however, as they are more difficult to control, may then be used to reveal reactions to stimuli that indicate a different mental state / belief than the one that the individual is willing to self-report.

An advantage of applying polygraph testing to the measurement of physiological reactions therefore is that it typically involves recording several ‘channels’ of physiological activity at once (Tredoux & Pooley, 2001), enabling a range of information (typically electrodermal activity (EDA), heart rate (HR) and respiratory rate) to be captured simultaneously. Additionally, Lykken (1981) states that physiological measurements using a polygraph are highly accurate provided that the apparatus is in good order and the polygraph operator follows generally accepted procedures. However, there are disadvantages of this method including; the expense of adequate equipment and frequent instrumentation problems with associated loss of data (Blascovich et al., 2011). There are often difficulties for example when interpreting the data output in relation to the effects of extraneous variables; such as cognition, temperature fluctuations etc. (Eisenberg & Fabes, 1990).

1.6.2. Physiological Measures of Emotion

As noted earlier, it has been suggested that self-report measures provide a useful method for assessing the subjective component of emotional responses. However, Fernández, Pascual, Soler, Elices, Portella, and Fernández-Abascal (2012) argue that using this method in isolation results in other components of emotional responses (i.e. physiological responses) being dismissed. Physiological reactions have been established as a fundamental component of emotion (Cacioppo & Gardner, 1999) therefore incorporating measures of physiological arousal can overcome the limitations of self-report methods as they allow for a more complex and

thorough objective understanding of emotional responding (Rosenthal et al., 2008). Baron-Cohen (2011) states that researchers within the field of empathy are resorting more frequently to including physiological measures of arousal.

When used within research assessing affect physiological measures have been argued to be superior to self-report measures in that they allow the researcher to measure change in affect over time while a stimulus is being displayed whereas self-report measures only capture summative emotions at a discrete point in time, usually after the stimulus has ceased (Lottridge, 2010). Feldstein and Gladstein (1980) when discussing the physiological measurement of empathy argue that physiological measurements may be a direct measure of empathic responsiveness such that the empathic state is evident before this is possible through verbal or cognitive means, as captured using self-report methods. Cacioppo and Gardener (1999) therefore conclude that although self-reports of affect are useful tools to elicit information about an individual's feeling state they nevertheless may be incomplete in the information they can capture from an individual.

Physiological indexes have been demonstrated to be valid markers of empathy-related emotions (Eisenberg & Fabes, 1990) and are therefore deemed useful in detecting emotions (Shi, Nguyen, Blitz, French, Fisk, et al., 2011). They have been used extensively in the field of emotion research to explore both general emotions (i.e. the identification of positive and negative affect; subtypes of distinct emotions i.e. anger, fear, happiness etc.; Kreibig 2010, for a review); as well as within clinical communication research (Finset, Stensrud, Holt, Verheul, & Bensing, 2011), and to explore continuous feedback of gaming experiences (Ben-Shakar, Bornstein, Hopfensitz, & van Winden, 2007), to name but a few research areas. A review of 134 publications reporting experimental investigations of the emotional effects on peripheral physiological responding in healthy individuals found that there is considerable evidence of autonomic nervous system response specificity in emotion when considering subtypes of distinct emotions (Kreibig, 2010).

Measurement of electrodermal activity (EDA) is generally considered to be one of the most sensitive physiological measures available (Gudjonsson, 1980), being the most widely studied response system in physiological research (Finset et al., 2011). EDA is measured by recording electrical activity in the skin. Eccrine glands located in the skin release sweat in response to autonomic nervous system arousal. As sweat increases so too does skin conductance. Increases in sweat may be gradual over time, referred to as tonic changes in skin conductance level (SCL). Or they may occur in immediate response to stimuli, resulting in sudden bursts of electrical activity, known as spontaneous skin conductance responses (SCR's). EDA is considered to provide a relatively direct and undiluted representation of sympathetic nervous system activity in response to stimuli when compared to other physiological measures (Finset et al., 2011) and is therefore a physiological marker of emotion commonly applied to emotional and attentional processes research. Hodgson and Rackman (1974) have suggested that the

concordance between physiological and subjective measures is affected by the intensity or level of emotional arousal.

Lang (1995) stated that SCL reflects emotional arousal and therefore argued that this can be used as a measure of emotional arousal. As SCL, or the converse skin resistance level (SRL), is not usually controlled by the individual then this provides a more reliable and objective measure of emotional arousal than self-reports of affect. When measuring EDA Dawson, Schell, and Fillion (2000) suggest that the palmar surfaces have been shown to be more responsive to significant or emotional stimuli than to thermal stimuli. In her review Kreibig (2010) reports that for EDA measures, SCL is the response variable most often reported, followed by skin conductance response rate and skin conductance response amplitude.

The cardiovascular response variable most often reported within Kreibig's (2010) review was identified as heart rate (HR), which Mandryk and Atkins (2007) state reflects emotional activity. Anastassiou-Hadjicharalambous and Warden (2007, p. 247) state that "HR increase...is an 'online' marker of vicarious affective arousal" (affective empathy) although Mauss and Robinson (2009) argue that there is no 'gold standard' of emotional responding. However, Thayer (1970) reported a combination of skin conductance and heart rate to be the best physiological composite measure, a combination that has been widely used in a number of studies to date, resulting in Kreibig (2010) terming them "convenience measures". Dawson et al. (2000) suggest the probable explanation for this popular combination is due to the utility of EDA as a general arousal indicator and HR for its potential differentiation of more subtle psychological states of interest to the researcher.

Skin conductance responses are generally considered to be components of the orienting response; behavioural and physiological responses elicited by either new or significant stimuli (Verschuere, Crombez, De Clercq, & Koster 2005). If stimuli are familiar or repeated, the orienting response is inhibited and habituation takes place. The orienting response may be confirmed by employing the simultaneous measure of HR to SCL; as HR deceleration in contrast to the increased EDA response (i.e. SCL), has been reported to be an index of the orienting response (Graham, 1979). Dawson et al. (2000) report that when monitoring an individual it is common for SCL to gradually decrease during rest periods, increase rapidly in response to novel stimuli and gradually decrease as the stimuli continues, a decrease also occurring with stimulus repetition. Overall, Mandryk and Atkins (2007) state that SCL correlates with arousal, presenting evidence mapping increased SCL to increased arousal. Taylor (1991) presented evidence showing that negative emotions were associated with stronger autonomic responses than positive emotions (although this study used a vastly disproportionate number of negative to positive emotions for comparison). More specifically, later research has shown different emotional expressions to produce different changes in EDA, for example irritation and anger have been found to increase EDA level, specifically HR and SCL, whereas positive emotions including happiness, contentment and affection have mixed results with reports of SCL activity including increased, decreased and unchanged levels (Kreibig, 2010).

Despite these findings Kreibig (2010, p. 410) concludes that, “there is no one-to-one relationship between emotion and changes in autonomic activation: feeling changes may occur without concomitant autonomic changes, just as autonomic changes may occur without concomitant feeling changes.” Moreover, there may be other factors affecting changes in EDA. For example research by Waid, Orne, and Wilson (1979) found that poorly socialised young adult males have been found to give smaller SCR’s to both physical and certain social stressors (i.e. deception / stressors) than their more highly socialised counterparts. Verschuere, Crombez, De Clercq, and Koster (2005) when investigating autonomic responding to concealed information, found that trying to conceal personal information elicits a predictable pattern of physiological reactions, specifically skin conductance responses become larger and heart rate declines when compared to control information. However, this is not so for antisocial offenders (those with high psychopathic traits) who instead display reduced EDA activity.

In summary, physiological measurements are a useful tool used within emotion and empathy research to gather objective measurements. However, as with self-report measures, they are not without their limitations. When applied to empathy research it has been argued that a combination of both self-report and non-self-report measures should be employed, as Eisenberg and Fabes (1990) argue that using one tool over another as a definitive index is not optimal. When assessing the relationship between empathy and pro-social behaviour in children Eisenberg and Fabes (1990) therefore employed a dual-method approach to data collection, combining both self-report and physiological measures; an approach they later applied to their empathy research with an adult population (also used by Ben-Shakhar et al. (2007) in their emotion research). This built upon work by Thayer (1970) who advocated for the method of combining physiological variables for correlation with self-reported data when assessing verbal reports of arousal and physiological variables. Gudjonsson (1980) has also demonstrated that combining the methods allows one to study the relationship between self-reported and physiological measures to generate hypotheses about what is being measured. Finally, by combining the methodologies and evaluating them simultaneously it is possible to ascertain the convergence among them (Fernandez et al., 2012), increasing the validity of the results. Despite this Fernández et al. (2012, p. 2) reported that “to date, studies that have examined the convergence among the measurement of different emotional elements have reported discrepant results”.

1.7. Combining Physiological and Self-Report Measures of Emotion

Emotional films have been used extensively in the study of emotion; reported to be one of the most popular and effective methods of emotion elicitation (Schaefer, Nils, Sanchez, & Philippot, 2010). Schaefer et al. (2010) suggest one main advantage of using emotional films compared to other methods within emotion research is the ability of the film exerts to elicit strong subjective and physiological changes in the viewer. Research has begun to move towards showing emotion-eliciting films to participants to induce an emotional response which is

subsequently assessed using a combination of self-report and physiological measures of emotion. However, Fernández et al. (2012) report that to date research using this convergent methodology in the field of emotion research has been mixed.

Fernández et al. (2012) investigated whether a set of emotion-eliciting films were able to induce measurable objective physiological responses, specifically if the films were able to induce physiological changes in SCL and HR. As a means to validate these findings they also employed subjective measures of emotional reactions in the form of a self-report measure, the Self Assessment Manikins (SAM; Bradley & Lang, 1994), arguing that a convergence between the two methods would demonstrate evidence for discrete emotions to be identified.

The SAM is a measure of emotional reactions encompassing three dimensions, affective valence, arousal and dominance. Results showed that in comparison to neutral films, SCL and HR were significantly increased after viewing fearful films; HR was also significantly increased after viewing anger films. HR was raised after viewing sad films although further analysis found this to be non-significant. They suggest that this may be so as these emotions create a more active behavioural response which would account for the greater increases in sympathetic activation and arousal in comparison to other emotions such as tenderness and sadness. Correlational analysis between variations in SCL and HR demonstrated that the measures were assessing similar physiological responses. Convergent analyses between the self-report and physiological measures showed a significant relationship between the arousal subscales of the SAM and both SCL and HR. However, no associations were found between the affective valence subscale of the SAM and either SCL or HR. In conclusion they suggested that viewing films that tap in to emotions with higher subjective arousal will result in the individual experiencing increased physiological activation. This provides support for Hodgson and Rachman's (1974) argument that the concordance between the two methods is affected by the intensity of emotional arousal (i.e. the more pertinent the stimuli to the observer the greater the concordance between the response systems). Taking this into account Gudjonsson (1980, p. 50) argued therefore that "it is perhaps not surprising to find a discrepancy between self-reported disturbance and electrodermal reactivity, especially at low levels of emotional arousal".

Anastassiou-Hadjicharalambous and Warden (2007) investigated the degree of convergence between three measures of affective empathy; physiological (change in HR), verbal self-report of emotion and facial expression when presenting emotionally evocative film clips to children. With regards to physiological responses, HR was analysed across baseline and during exposure to four evocative film clips containing the emotions; sadness, fear, anger and happiness/surprise. Analysis revealed a significant effect of film clip such that HR was significantly higher relative to baseline in exposure to the fear clip, followed by anger, sadness and happiness/surprise. However, convergence analysis between verbal self-report of the child's emotional reaction to the film clip and change in HR only occurred at chance levels. They suggest this lack of convergence may be accounted for by the limitations associated with verbal self-report of emotions. They argue that children may struggle to identify and describe

their feelings with accuracy thus there may be a discrepancy between what children in the study experienced affectively and how they interpreted and reported such experiences; an explanation supported by Eisenberg and Fabes (1990) when discussing the limitations of using self-report measures of empathy with offenders. They concluded therefore that if one is interested in identifying the conscious emotional experience only, then self-report measures are recommended. However, if the focus is to accurately and reliably measure the presence and degree of vicarious affective arousal (affective empathy) then physiological indices, although lacking in clarity regarding valence and specific emotion, are recommended.

Evidence presented so far of physiological reactions to emotion-inducing stimuli has typically suggested that where an emotion-inducing stimulus is presented an individual will have a subsequent emotional response to this, typically an increase in SCL and HR. However, there are some groups who go against this trend, for example poorly socialised young adult males and antisocial offenders (high in psychopathy traits) who instead show reduced EDA activation when responding to fear inducing stimuli. Evidence has also been presented suggesting that where an individual deems a stimulus to be highly arousing, established via self-reports of subjective emotional state, a concurrent physiological response would be expected, again increased SCL and HR. However, it has been argued that how one copes with stress in situations of increased emotional arousal, including how one perceives oneself to cope and how one subsequently responds to this, can also cause a discrepancy between self-report and physiological measures of emotional disturbance.

Two groups have been identified which show how people react to emotional stimuli, namely 'repressors' and 'sensitisers'. 'Repressors' are described as those who report low subjective emotional disturbance to emotion-inducing stimuli yet have relatively strong EDA reactions. Characteristically they have been described as individuals who are preoccupied with being in control of their emotions, presenting a somewhat calm, rational demeanour. Conversely 'sensitisers' are those who tend to over amplify their subjective emotional disturbance to emotion-inducing stimuli yet who have relatively low EDA reactions, described as individuals who characteristically amplify their view of themselves as anxious (Weinberger, Schwartz, & Davidson, 1979). Gudjonsson (1980) reported that in general these groups are highly inaccurate in judging their emotional disturbance in relation to emotionally loaded questions. He also demonstrated that it is possible to identify these groups through their personality and defensiveness traits; repressors have high defensiveness and low trait anxiety whereas sensitisers have low defensiveness and high trait anxiety. Furthermore he reported that the inaccuracy of self-reported emotional disturbance is significantly related to defensiveness and trait anxiety.

Vlahou, Vanman, and Morris (2011) argue that people use two strategies to regulate their emotional response to emotion-eliciting stimuli termed 'emotional reappraisal' and 'emotional suppression'. Emotional reappraisal represents one's ability to alter the emotional impact of a stimulus quickly, thus preventing a true emotional response; this involves a cognitive change in

how one appraises the stimuli. Emotional suppression however represents one's efforts to control their emotional response for example, deliberate inhibition of emotionally expressive behaviour e.g. crying, verbal reports to the contrary, even though one is feeling emotionally aroused. Emotional suppression therefore is characterised by self-reports of decreased negative feelings but increased physiological activation, compared to emotional reappraisal. Boden and Dale (2001) investigated this using a film clip depicting dying baby parrots used to elicit unpleasant emotions in the observer in comparison to an emotionally neutral film clip. They found that repressors engaging in emotional suppression reported less negative affect compared to sensitizers when exposed to the unpleasant clip; no group differences in affect were found for the neutral clip. Sparks, Pellechia, and Irvine (1999) used frightening film clips to demonstrate that although repressors reported low levels of negative affect after viewing the clips, in contrast their physiological arousal was increased to greater levels than non-repressors. Finally, Brosschot, and Janssen (1998) demonstrated discordance between rating of tenseness and physiological arousal which was greater for repressors in comparison to non-repressors when exposing participants to fear-eliciting films which were unrelated to violence. Finset et al. (2010) provide support for the process involved in emotional suppression when discussing emotional regulation research. They suggest, after finding increases in physiological responses despite an individual successfully regulating his/her subjective emotional state, that physiological responses may therefore be considered as a marker of inhibited subjective expression of emotions.

Combining these findings it is suggested that repressors would typically employ emotional reappraisal or emotional suppression strategies. Research has argued that sex offenders engage in cognitive reappraisal to justify their actions for example, reports from rapists often state that they believe the victim enjoyed the experience (Fisher, Beech, & Browne, 1999). Similarly where offenders accept their crimes but are overcome with feelings of shame or guilt associated with this they may engage in emotional reappraisal. However, it is suggested that when viewing an emotion-eliciting film that is particularly emotionally salient to the observer, the observer would respond using emotional suppression techniques. For example, in an experimental context the observer may attempt to control their emotional response by subjectively denying associated negative affect i.e. upset, distress, yet nevertheless display their 'true' distress through their physiological reactions. This would create discordance between their subjective and objective measurements of emotional reaction. With regards to assessing for cognitive and affective empathy this may also demonstrate discrepancies for example, in response to self-reported empathy after viewing an emotion-eliciting film; if engaging in emotional suppression one may self-report low levels of distress yet in contrast may accurately identify that a victim would feel distress in this situation.

Overall, it has been suggested that measuring empathy with offenders has typically relied on self-report measures. However, there are numerous limitations associated with this method in isolation. Physiological indexes, specifically SCL and HR have been demonstrated to be valid markers of empathy-related emotions (Eisenberg & Fabes, 1990), although again this method

also has associated limitations. Consequently, it has been argued that using one method over the other as a definitive index of studying emotion is not optimal. A combination of both self-report and non-self-report measures have therefore been advocated for use within the field of empathy research, as the self-report measure provides a useful method for assessing the subjective component of emotional responses whereas measures of physiological arousal allow for a more complex and thorough objective understanding of emotional responding (Rosenthal et al., 2008). Convergent analysis of the data generated by using both methods will have the benefit of increasing the validity of this data. Emotion-eliciting films have proved to be the most popular and effective format to assist in the generation and measurement of emotion in participants (Schaefer et al., 2010).

Research with sex offenders has increasingly employed the use of polygraph testing in the clinical setting. This has been used to measure offenders' physiological responses to stimuli over time. It also has the advantage of eliciting information from the offender that is otherwise difficult to attain (i.e. the offender may wish to be untruthful in his responses to self-report measures of empathy as he wishes to be seen in a socially desirable way, or he may not be aware of his emotions to be able to accurately identify these). To date, research using combined self-report and physiological methods within emotion research has had mixed results, specifically with regards to identifying discrete emotions. However, it has been argued that physiological measures reflect one's level of alertness and emotional intensity generated from emotion-inducing stimuli (Fernández et al., 2012), with greater emotional stimuli creating greater concordance between the measures and low emotional stimuli frequently presenting a discordance. One must also be aware of a number of factors that may influence the discordance between self-report and physiological measures of emotion including; personality attributes of the individual, including if an individual may be a 'repressor' or a 'sensitiser'. The VERA(ii) is a new measure of victim empathy that requires validation although there are numerous limitations associated with the validation methods currently employed (i.e. when using existing self-report victim empathy measures). As the VERA(ii) consists of emotion-eliciting video interviews and contains both cognitive and affective self-report measures of empathy it is therefore suggested that employing additional physiological measurements will assist with the validation of the tool.

2. RATIONALE

It has been argued that offenders have deficits in victim empathy, which results in them behaving in undesirable ways (i.e. engaging in anti-social behaviour, particularly violent and sexual offences). For this reason developing empathy has become a central component of specialised cognitive behavioural treatment programmes, including relapse prevention. An additional benefit derived from appropriate and effective treatment programmes being the important economic benefits also. It is therefore important to be able to evaluate the effectiveness of these interventions. Research has concluded that measures of victim empathy are of greater clinical utility for this purpose compared to measures of general empathy. The most common victim empathy measures are self-report questionnaires of which numerous limitations have been highlighted including; the measures often differ widely in their definition of empathy resulting in different components of the empathy construct being measured (i.e. cognitive empathy and / or affective empathy or general empathy); consequently the psychometric properties of the measures are inconsistent. The purpose of these measures are generally transparent making them susceptible to response bias or 'faking good' by the offending population with whom they are used. In light of these criticisms it has been argued that better measures of empathy are needed. Young et al. (2008) developed the VERA in response to this.

The validity of the VERA has been assessed with both a forensic sample and an adult male community sample and has demonstrated some limited evidence. Following participant feedback of how to increase the validity of the tool a second prototype VERA(ii) was developed, which also needs validating. Acknowledging the limitations of existing measures from which to validate a new tool, it has been suggested that additional validation methods need to be employed alongside existing self-report methods, for example physiological measures. Research with offenders has typically employed physiological measures using polygraph testing to assist with identifying deception. Physiological measures have also been used extensively to assess emotion, based upon the analysis of an individual's EDA. For example, SCL and HR have been shown to monitor emotional arousal; increases in SCL with simultaneous decreases in HR being shown to occur when an individual experiences negative emotions, as well as when trying to regulate emotions. Employing physiological methods in this way in combination to self-report measures will allow one to assess the emotional arousal of the respondent to detect if he is indeed experiencing an empathic response or not as he reports on the self-report part of the victim empathy measure (i.e. VERA(ii), which will increase the validity of the tool). However, there is evidence that physiological reactions can be affected by other variables, for example personality (i.e. antisocial offenders; those with high psychopathy traits); 'theory of mind deficits' (i.e. persons on the autistic spectrum and persons with a diagnosis of schizophrenia) have been shown to have deficits in empathy such they do not produce increases in arousal even when being deceptive; as well as intelligence (i.e. low intelligence has been associated with low

empathy). Therefore, taking this into consideration, the research aims and hypotheses of the current study follow.

3. AIMS AND HYPOTHESES

The current study was conducted with the primary aim of validating the VERA(ii) with an adult male community sample using an existing self-report measure of empathy and physiological measures of arousal. Co-variables including social comprehension, intelligence, neuroticism, anti-social personality traits and violent attitudes and were also included as evidence has shown these to have an association with empathy.

The hypotheses generated as a result of these aims are:

3.1. Hypotheses

1. (a) Both total cognitive and affective empathy constructs will demonstrate satisfactory internal consistency overall for VERA(ii) and (b) across all vignettes of the VERA(ii)

2. VERA(ii) total cognitive and affective empathy self-report scores will be positively correlated with an existing global empathy self-report measure

Higher scores on the VERA(ii) self-report subscales will be associated with higher scores on the IVE-Empathy subscale

3. VERA(ii) total cognitive and affective empathy self-report scores will be negatively correlated with SRL and positively correlated with HR, during stimulus presentation

Higher scores on the VERA(ii) self-report subscales will be associated with greater change in mean SRL and HR from baseline to stimulus presentation; where mean SRL will decrease and mean HR will increase

4. VERA(ii) total cognitive and affective empathy self-report scores will be negatively correlated with (a) antisocial personality traits and (b) violent attitudes and positively correlated with (c) social comprehension and (d) neuroticism and, (e) IQ

(a) Higher scores on the VERA(ii) self-report subscales will be associated with lower scores on the EPQ-R Psychoticism subscale

(b) Higher scores on the VERA(ii) self-report subscales will be associated with lower total scores on the MVQ Machismo and Acceptance sub-scales

(c) Higher scores on the VERA(ii) self-report subscales will be associated with higher scores on the Story Comprehension Test

(d) Higher scores on the VERA(ii) self-report subscales will be associated with higher scores on the EPQ-R Neuroticism subscale

(e) Higher scores on the VERA(ii) self-report subscales will be associated with greater estimated Full-Scale IQ

4. METHOD

4.1. Design

The study was a cross-sectional within-subjects design.

4.2. Power Analysis

The required sample size necessary to detect a significant effect in the data was based on data from Terry, Gudjonsson, and Young (2009) who assessed the psychometric properties and validity of the audio format of VERA. Using 100 normal participants they reported a positive correlation of 0.33 between anti-social personality traits (as measured by the Psychoticism scale of the Eysenck Personality Questionnaire – Revised (EPQ-R; Eysenck, Eysenck, & Barrett, 1985) and cognitive empathy. Power analysis using nQuery Advisor 4.0 indicated that a sample of 50 participants would be needed to achieve this level of association with 80% power at a 0.05 one-tailed significance level.

4.3. Participants

51 healthy adult males from the general population were recruited, aged 20 to 55 years. Their mean age was 29.2 years (SD = 9.3). Over half of participants (58.8%) identified their ethnic origin as White British, 9.8% were White European, 7.8% were Asian and 3.9% were Black African. 19.6% stated their ethnic origin as Other.

4.4. Inclusion / Exclusion Criteria

Inclusion criteria stipulated that participants should be English speaking males, aged between 20-60. Exclusion criteria included; those who have convictions for serious offences and/or history of severe mental illness in addition to those who have a learning disability, a history of serious head injury, or a neurological condition that significantly affects mental functioning.

4.5. Recruitment

Participants were accessed and recruited from two locations.

1. Volunteers registered with the Mindsearch database at the Institute of Psychiatry (IoP) to participate in research studies
2. Staff and students at Kings College London University (KCL)

Participants from Mindsearch and KCL were initially contacted via a circular e-mail. The e-mail contained brief details of the study including what was involved, the inclusion and exclusion criteria and that they would receive a payment of £10 for their time should they wish to

participate. Those interested in taking part were advised to contact the researcher for more information. A detailed information sheet was then forwarded to those who responded, see Appendix 1. Participants contacted the researcher again if their interest remained and they wished to opt-in to the study. A suitable time was then arranged at their convenience to come to the IoP to take part. All participants were paid £10 for their time.

Ethical approval was granted by the North London Research Ethics Committee 3 (Harrow), study reference number 11/LO/0314. Ethical approval was granted on two occasions; the original proposal, and following the introduction of a substantial amendment to incorporate an IQ measure into the study, see appendices 2 and 3 respectively. Subsequently, Research and Development approval was granted from West London Mental Health Trust, study reference number YOUSW1101.

4.6. Apparatus and Software

VERA was displayed using a HP Compaq 6735s laptop, running Windows Vista. The Lafayette LX4000 Polygraph System running LXSoftware, displayed on a Dell Latitude c800 laptop, was used for physiological recordings. This instrument was chosen for use within the current study as it is typically used with offenders within forensic settings. Only the skin resistance and heart rate channels of the Lafayette LX4000 Polygraph System were utilised by the current study.

4.7. Materials

4.7.1. Victim Empathy Response Assessment (ii) (developed by Young et al.)

VERA(ii) (developed by Young et al.) was designed to provide an assessment of cognitive and affective empathy in forensic settings. It is a programme consisting of 5 video vignettes of female 'victims' being interviewed describing their experiences of different incidents including; arson, car accident, assault, child sexual abuse and rape. Following each interview participants are asked to rate 13 items using a five-point Likert scale (0 = 'Not at all'; 4 = 'Very much so'), how they felt while watching the interview (affective empathy; see Appendix 4) and how they believe the victim felt about her experience (cognitive empathy; see Appendix 5). Total scores for the scales are calculated from summing the 9 'distress' items, as highlighted by Young et al. (2008), independently for both cognitive and affective empathy for each participant across each vignette. Total cognitive and affective empathy scores are calculated for each participant by summing the total scores for all vignettes.

VERA(ii) was developed following feedback from an initial audio format (VERA) that consisted of 5 tape-recorded staged 'radio broadcasts' of victims describing their experiences of the incidents stated above, where it was highlighted that the audio format seemed artificial. It was therefore deduced that the reliability of the tool could possibly be improved through the introduction of a visual format; VERA(ii) has therefore been developed.

Initial validation of the VERA by Young et al. (2008) using a forensic sample demonstrated good internal reliability across all vignettes (Cronbach's α ranging from 0.85 to 0.94) also replicated with a community sample (Cronbach's α ranging from 0.90 to 0.94; Terry et al., 2009). There is some evidence for construct validity, for example Terry et al. (2009) found significant positive correlations between the VERA and the Empathy Subscale of the IVE (Eysenck & Eysenck, 1978).

4.7.2. Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999)

The two subtest form of the WASI (Vocabulary and Matrix Reasoning) was used to provide an estimate of Full-scale intelligence, in order to investigate its relationship with empathy, and if significant control for the potentially confounding effect between empathy and the other measures. The two subtest of the WASI has good internal consistency (Cronbach's α = 0.96) and good test-retest reliability (r = 0.88) (Wechsler, 1999).

4.7.3. Eysenck Personality Questionnaire – Revised; Short Scale (EPQ-R Short Scale; Eysenck, Eysenck & Barrett, 1985)

The EPQ-R Short Scale is a 48 item yes/no self-report inventory which measures three key dimensions of personality; Extraversion (E), Neuroticism (N) and Psychoticism (P), a Lie scale (L) is also included. Each scale consists of 12 items. The scales have satisfactory reliabilities (E = 0.88; N = 0.84; P = 0.62; L = 0.77); reported here for males. The EPQ-R Short Scale has been widely used, including by Alexio and Norris (2000) when assessing personality and moral reasoning in young offenders.

4.7.4. Eysenck Impulsiveness, Venturesomeness and Empathy Questionnaire (IVE; Eysenck, Pearson, Easting, & Allsopp, 1985)

The IVE is a 54 item yes/no self-report inventory which measures three personality traits; Impulsiveness (Imp), Venturesomeness (Vent) and Empathy (Emp). Overall the scales have satisfactory reliabilities (Imp = 0.84; Vent = 0.85; Emp = 0.69) and good test-retest reliabilities (Imp = 0.78; Vent = 0.85; Emp = 0.77); reported here for males (Eysenck et al., 1985). When correlated with the EPQ, Eysenck and Eysenck (1975) reported Empathy to be positively related to Neuroticism but negatively related to Psychoticism (Eysenck & Eysenck, 1991).

4.7.5. Maudsley Violence Questionnaire (MVQ; Walker, 2005)

The MVQ is a 56 item true/false self-report inventory which measures two violent cognitive styles: (i) 'Machismo' (42 items) is a construct around believing that it is an embarrassment to back down from violence, justifying violence in response to a perceived threat and viewing aggression as a strength, and (ii) 'Acceptance of Violence' (14 items), a construct around being generally accepting of violence such that it can be enjoyable in certain contexts (Walker, 2005).

The MVQ was initially validated using a 'normal' student adolescent population and reported the two constructs to have adequate internal consistency (Machismo: Cronbach's $\alpha = 0.91$, Acceptance: Cronbach's $\alpha = 0.74$; reported here for males (Walker, 2005). Further validation of the two constructs has been assessed with this population by evaluating the relationship between personality and self-reported offending (Walker & Gudjonsson, 2006). For males machismo was found to be the strongest predictor of self-reported violence ($r = 0.50$), and EPQ-R psychoticism was the strongest predictor of self-reported non-violent offending ($r = 0.39$).

Using an adult forensic clinical population Warnock-Parkes, Gudjonsson, & Walker (2007) investigated the relationship between the two constructs and objective measures of violence. In accordance with predictions they reported machismo to be significantly related to number of past criminal convictions ($r = 0.25$) and institutional violence ($r = 0.36$) and acceptance to be significantly related to institutional violence ($r = 0.29$). At present there are no total MVQ score norms available.

4.7.6. Story Comprehension Task (SCT; Channon & Crawford, 2000)

The SCT is a measure of social comprehension / theory of mind. It consists of 12 written vignettes each describing a brief story. Participants are requested to read the story (or have this read to them if they have literacy difficulties or visual impairments) after which they are asked to explain the reason behind the main characters speech or actions. The stories include examples of sarcasm, pretence and misunderstanding, among others. They are constructed so that they cannot be understood using a simple literal interpretation, making it necessary for the participant to generate a non-literal interpretation. Answers are deemed correct if the participant has been judged to have interpreted the words or actions of the main character accurately.

Findings reported that participants with left anterior lesions commonly failed to make non-literal inferences on the story comprehension test relative to healthy participants. As brain imaging studies commonly indicate areas of the left frontal lobe to be activated in theory of mind tasks they suggest therefore that failure to make non-literal inferences on the story comprehension task is indicative of theory of mind deficits.

4.8. Procedure

Participants were tested individually at the IoP. When the participant arrived, he was seated in a comfortable, adjustable chair. The researcher explained the procedure after which the participant completed an informed consent form and a brief health questionnaire. Participants were not told that the victims being interviewed in the vignettes were actresses. If the participant asked, the researcher responded that they did not know and so were unable to clarify this. Physiological sensors were then attached to the participant's non-dominant arm / hand; this was established by the researcher asking the participant which hand he would use to

control the mouse when selecting his answers from the electronic questionnaire. Electrodermal activity (SRL) was recorded by attaching disposable pre-gelled electrodes (Ag/AgCl, contact area 1cm diameter) to the palmar surface of the first and third fingers of the non-dominant hand. An electrode cable linking to the data acquisition system of the Lafayette LX4000 polygraph was attached to these via snap connectors. Heart rate was measured using a standard inflatable cuff placed on the participant's non-dominant arm. Once the recording equipment was in place the participant was asked to adjust the chair so that he was sitting comfortably with his legs uncrossed, and feet flat on the floor facing a laptop that had been placed on a table directly in front of him. He was asked to rest in the chair with his hands facing palm down, placing either his arms on the arm rests of the chair or on the table either side of the laptop; whichever he felt was the most comfortable position. A baseline recording of SRL and HR was then taken for one minute. During this time the participant was instructed to sit as previously advised and to remain as still as possible while looking straight ahead to the laptop, which was displaying the desktop screen.

Once the baseline recording was complete, the VERA(ii) was presented. This consisted initially of an introduction to the measure where a recorded male voice read out written instructions that were displayed on the screen in front of the participant. The participant was then instructed to complete examples, in the format of the questionnaire items to follow, to familiarise himself with the structure of the VERA(ii) questionnaires. Once the introduction to VERA(ii) was complete the researcher inflated the HR cuff, instructed the participant to sit as he did during the baseline recording and then press begin on the laptop. The participant then watched the first of the five VERA(ii) vignettes. Each vignette was preceded by a recorded male voice stating which vignette was beginning for example, "Interview 1" followed by the corresponding video interview / vignette (i.e. vignette 1). Upon hearing this prompt the researcher simultaneously began the physiological recording.

If the participant displayed a clear movement during this time causing a fluctuation to appear on the physiological recording that was clearly the result of the voluntary movement and not a consequence of an involuntary physiological reaction to the visual stimuli, these were termed 'artifacts'. A note was made by the researcher as to the exact time of the artifact and a brief description of the artifact was also recorded.

The physiological recording was terminated after each vignette was finished; see Appendix 6 for an example of physiological recording output. At this time the HR cuff was deflated and the participant was requested to complete the two electronic VERA(ii) questionnaires; these were in the order of the affective empathy questionnaire, followed by the cognitive empathy questionnaire. This was in keeping with the order used by both Young et al. (2008) and Terry et al. (2009). Once the questionnaires were completed the procedure was repeated for the remaining four vignettes. The vignettes were displayed in a pre-set format as received from the manufacturer commissioned to design the VERA(ii) database in the order: Arson; Car Accident; Assault; Child Sexual Abuse and Rape. Once the VERA(ii) had been completed the

physiological recording equipment was removed from the participant to allow him to complete the remaining part of the study.

The remaining part of the study consisted of the participant completing the Vocabulary and Matrix Reasoning sub-tests of the WASI, administered by the researcher. Following this the participant completed the self-report measures in the following order: EPQ-R, IVE, and MVQ before the SCT was administered by the researcher. At the end of the session the participant was fully debriefed by the researcher, including confirming that the victims were in fact actresses. Any questions he had were answered by the researcher before receiving his £10 payment.

4.9. Statistical Analysis

4.9.1. Physiological Data

Changes in tonic EDA were calculated using the method employed by Finset et al. (2011); however by subtracting mean levels of skin resistance during baseline from mean levels of skin resistance (as opposed to skin conductance) during each vignette for each participant, and will be referred to as skin resistance level (SRL). This method was also applied to the HR output to calculate mean change in HR.

Where artifacts were present in the data these were removed by the researcher. Artifact removal was deemed necessary as it has been argued that the presence of artifacts (i.e. unwanted variation in the measured signal as a result of sources external to those being investigated) in physiological recordings seriously degrades the utility of the measurements taken (Sweeney, Ward, & McLoone, 2012). The procedure for removing artifacts included identifying, from the information previously recorded by the researcher, where the artifact lay on the electronic graph output. The researcher then removed this section from the output for each vignette, including the baseline recording if artifacts were present here also. The mean, minimum and maximum scores of the remaining sections of the output for each interview were recorded, these were then used to calculate an overall mean, minimum and maximum score for each interview, including baseline if necessary. Where no artifacts were recorded the original mean, minimum and maximum scores of each interview were used.

The mean, minimum and maximum SRL and HR scores were calculated for each set of participant data. Each set of data contained six items including: baseline and vignettes 1-5. Approximately 1/5th (19.6%) of participant's data sets contained artifacts, with the number of artifacts per set ranging from 1 – 5. Therefore of the 306 total items within these sets, 101 items (33%) contained artifacts; with the number of artifacts per item ranging from 1- 5.

4.9.1.1. Inter-rater Reliability

As artifacts were removed manually, based on the judgement of the researcher (being guided by the artifact data recorded previously), reliability of this procedure was deemed necessary. A second researcher assisting with the study therefore independently conducted the procedure of artifact removal from 10% of the sample ($n=5$). Following discussion between raters, and due to the length of time taken to complete the procedure, 10% of the sample was deemed to be sufficient to assess for reliability. Inter-rater reliability using the Intraclass Correlation Coefficient (ICC) Two-way Random method was then calculated for total agreement between raters for both SRL and HR mean, minimum and maximum scores. This method was chosen as it allows one to determine how accurately raters are returning the same score for an item. A total of 90 ratings per SRL and HR were compared (30 items containing 3 ratings; mean, minimum and maximum scores, for each SCL and HR). An almost perfect level of agreement was present for both SRL ratings ($ICC = 1.00$, $p < 0.01$) and HR ($ICC = 1.00$, $p < 0.01$). ICC was also 1 for each SRL and HR item when assessed individually. Using these results, the data was therefore deemed reliable for further analysis.

4.9.2. Data Treatment

Prior to inferential analysis, all data (self-report measures and physiological) was examined to assess if the assumptions of normality were met, as necessary to conduct parametric statistical tests. Data was inspected visually using histograms and quantile-quantile plots (Q-Q plots) to assess for normality of the distribution scores. The visual interpretations were confirmed by calculating z-scores for skewness and kurtosis (which compare the symmetry and cluster of the data against that of a normal distribution). Finally Kolmogorov-Smirnov and Shapiro-Wilks statistical tests were also performed.

Combined, the methods indicated that the majority of the data did not fulfil the assumptions of normality necessary to conducted parametric tests. Transformations of the data including: log, square root and reciprocal transformations were therefore applied to determine if the distributional problems of the data could be corrected to that of a normal distribution. However, this was unsuccessful for the majority of the data, specifically the cognitive empathy data. Non-parametric tests using Spearman's Rho were therefore chosen to investigate the correlations between variables. This was also in keeping with the analysis employed by Terry et al. (2009) in their study validating the original VERA.

Where differences were found between groups (i.e. cognitive and affective empathy, and between physiological measures) post-hoc investigations were conducted employing Wilcoxon signed-rank tests. The Wilcoxon-signed rank test was chosen as only some of the data was not normally distributed making it necessary to employ this non-parametric test.

A minimum significance level of $p < 0.05$ was employed to the interpretation of all analyses.

Cohen's criteria was chosen as an effect size measure to assist with the interpretation of any correlations found, where Cohen (1988, 1992) states: $r = .10$ (small); $r = 0.3$ (medium); and $r = 0.5$ (large) effect size.

5. RESULTS

As outlined in the Method, data from 51 participants was collected. The following section will first outline the descriptive results of the self-report and physiological data before consideration of the main hypotheses.

5.1. Descriptive Statistics:

5.1.1. VERA(ii)

The mean, standard deviation (SD), median and range of scores for total cognitive and affective empathy are presented in Table 5.1.

Table 5.1. Mean, Standard Deviation, Median and Range of Scores for Total Cognitive and Affective Empathy (n=51)

VERA(ii) Subscale	Mean	SD	Median	Range
Cognitive Empathy	139.82	25.58	144	0-170
Affective Empathy	86.76	35.44	90	6-161

On average total cognitive empathy scores were greater than total affective empathy scores, with cognitive empathy scores showing less variance from the mean compared to affective empathy scores. The range of scores for both constructs was large with some participants scoring extremely low on both constructs, in-particular for cognitive empathy where no empathy was reported, yet others achieved almost the maximum score (i.e. 180).

Wilcoxon signed-rank test was employed to establish if the difference between the constructs was significant. Table 5.2. confirms that this difference was significant. The effect size calculated using the formula for non-parametric tests recommended by Field (2005) is large and is shown in the table.

Table 5.2. Wilcoxon test results for the difference between Total Cognitive and Affective Empathy (n=51)

z score	Effect size
-6.093**	.85

** $p < 0.001$

Furthermore, Spearman's Rho Correlation Coefficient was calculated to assess if a relationship existed between the cognitive and affective empathy construct. Table 5.3. confirms that there was a significant positive correlation between the constructs with a medium effect size.

Table 5.3. Spearman's Rho Correlation of Total Cognitive and Affective Empathy

Total Affective Empathy
Total Cognitive Empathy .49

** $p < 0.01$

5.1.2. Self-Report Measures

The mean, SD, median and range of scores for self-report measures and IQ, are presented in Table 5.4.

Table 5.4. Mean, SD, Median and Range of Scores for Self-Report Measures, and IQ (n=51)

Measure	Mean	SD	Median	Range
IVE-Empathy	12.75	3.47	13.00	5-20
EPQ-R Short Scale Psychoticism	3.01	1.82	3.00	0-7
EPQ-R Short Scale Neuroticism	3.82	3.33	2.00	0-12
EPQ-R Short Scale Lie	3.75	3.58	3.00	0-13
MVQ Machismo	3.71	4.91	2.00	0-27
MVQ Acceptance	8.06	3.31	8.00	1-13
SCT	10.75	1.44	11.00	5-12
WASI Predicted Full-Scale IQ	122.52	9.60	124.00	87-140

The main score to note in Table 5.4. is the high IQ among this sample, falling approximately 1.5 standard deviation's above the mean. The average estimated full-scale IQ of the sample was therefore 'high average'. IQ ranged from 'low average' to 'high average' however the medium IQ score was within the 'high average' range. All other measures were similar, falling within 1 standard deviation of their respective 'normative' mean values previously published.

5.1.3. Physiological Measures

The mean, SD, median and range of physiological measures, are presented in Table 5.5.

The range of scores for both mean baseline SRL and mean total SRL were large with the median values being considerably lower than the maximum values recorded, as presented in Table 5.5. Overall there was a decrease in mean baseline SRL compared to mean total SRL as shown by the change in total mean SRL score.

Table 5.5. Mean, SD, Median and Range of Physiological Measures (n=51)

Physiological Measurements	Mean	SD	Median	Range
Mean Baseline SRL (ohms)	131.51	116.25	104.00	48.00-653.00
Mean Total SRL (ohms)	119.22	97.11	96.60	59.60-555.40
Change in Total Mean SRL (ohms)	-12.29	27.47	-6.40	-97.60-45.60
Mean Baseline HR (bpm)	65.78	12.33	65.00	36.00-96.00
Mean Total HR (bpm)	66.29	11.74	66.8	36.80-98.60
Change in Total Mean HR (bpm)	0.51	5.24	0.40	-10.20-15.00

A Wilcoxon signed-rank test was employed to establish if the difference between mean baseline SRL and mean total SRL was significant. Table 5.6. confirms that there was a significant difference. The effect size calculated using the formula for non-parametric tests recommended by Field (2005) is medium and is displayed in the table.

Table 5.6. Wilcoxon test results for the difference between Mean Baseline and Total Mean for both SRL and HR (n=51)

	z score	Effect size
SRL	-2.943*	.41
HR	0.560	.08

* $p < 0.01$

The average mean baseline and mean total HR scores were situated around the medium score of the data. There was a slight increase in mean baseline HR compared to mean total HR as shown by the change in total mean HR, as shown in Table 5.5. A Wilcoxon signed-ranks test found this difference to be non-significant, as shown in Table 5.6.

5.2. Hypotheses

5.2.1. Hypotheses 1(a) and 1(b)

Cronbach's α for VERA(ii) cognitive and affective empathy sub-scales for each vignette and total vignettes are presented in Table 5.7.

Table 5.7. Reliability (Cronbach's α) for VERA(ii) Cognitive and Affective Empathy Subscales for each vignette and total vignettes (n=51)

	Cognitive Empathy	Affective Empathy
Interview 1 – Arson	.84	.90
Interview 2 – Car Accident	.80	.87
Interview 3 – Assault	.83	.91
Interview 4 – Child Sexual Abuse	.92	.93
Interview 5 – Rape	.91	.94
Total Score	.87	.94

Overall, both cognitive and affective empathy subscales achieved high reliability:

Cognitive empathy scale; $\alpha = 0.87$; Affective empathy scale; $\alpha = 0.94$

Reliability analysis using Cronbach's α demonstrated high internal consistency across all interviews for cognitive empathy (range 0.80 – 0.92), and affective empathy (range 0.87 - 0.94), see Table 5.7.

Table 5.8. Item Analysis for VERA(ii) Cognitive and Affective Empathy Subscales (n=51)

	Cognitive Empathy		Affective Empathy	
	Item-Total Correlation	Cronbach's α if item deleted	Item-Total Correlation	Cronbach's α if item deleted
Interview 1 - Arson	.76	.84	.82	.94
Interview 2 – Car Accident	.76	.83	.83	.94
Interview 3 - Assault	.77	.83	.86	.93
Interview 4 – Child Sexual Abuse	.50	.91	.89	.92
Interview 5 - Rape	.83	.82	.88	.93

Item-total correlations were calculated to demonstrate how the subscales functioned within the measure, as shown in Table 5.8. Norman and Streiner (2008) state that item-total correlations should be greater than $r = .3$ (preferably $r = .4$) but not greater than $r = .8$.

As can be seen in Table 5.8., all items with the exception of Interview 5 – Rape, are within the acceptable range for cognitive empathy. However, only the removal of Interview 4 – Child Sexual Abuse would increase the overall reliability of the scale. Within affective empathy all items are above the upper range ($r = .8$), however removal of any of these items would not increase the overall reliability of the scale.

5.2.2. Hypotheses 2 and 4

Spearman's Rho Correlation Coefficients were calculated to assess the relationship between VERA(ii) total cognitive and affective empathy sub-scales and self-report measure sub-scales, and IQ; displayed in Table 5.9.

Table 5.9. Spearman's Rho Correlations of Cognitive and Affective Empathy with Self-Report Measures, and IQ (n=51)

Self-Report Measures and IQ	Cognitive Empathy	Affective Empathy
IVE Empathy	.12	.21
EPQ-R Psychoticism	-.11	-.22
EPQ-R Neuroticism	-.04	-.06
EPQ-R Lie	.09	.14
MVQ Machismo	-.27*	-.02
MVQ Acceptance	-.14	-.15
SCT	.30*	.01
FSIQ	.23	-.06

* $p < 0.05$

A significant negative association was found between cognitive empathy and violent attitudes with a small effect size. A significant positive correlation was found between cognitive empathy and social comprehension with a medium effect size. No other significant associations were found between total cognitive or affective empathy and the remaining self-report measures.

5.2.3. Hypothesis 3

Spearman's Rho Correlation Coefficients were calculated to assess the relationship between VERA(ii) total cognitive and affective empathy sub-scales and physiological measures; displayed in Table 5.10.

Table 5.10. Spearman's Rho Correlations of Cognitive and Affective Empathy with Physiological Measures (n=51)

Physiological Measures	Cognitive Empathy	Affective Empathy
Total Mean SRL	.00	.15
Total Mean Change SRL	.07	-.20
Total Mean HR	-.11	-.17
Total Mean Change HR	-.13	-.12

There were no significant correlations found between either cognitive or affective empathy and physiological measures.

Spearman's Rho Correlation Coefficients were calculated to assess the relationship between mean total SRL and mean total HR and are displayed in Table 5.11.

Table 5.11. Spearman's Rho Correlation between Mean Total SRL and Mean Total HR (n=51)

	Mean Total HR
Mean Total SRL	-.20

No significant association was found between mean total SRL and mean total HR.

Spearman's Rho Correlation Coefficients were calculated to assess the relationship between total mean change SRL and total mean change HR and are displayed in Table 5.12.

Table 5.12. Spearman's Rho Correlation between Total Mean Change SRL and Total Mean Change HR (n=51)

	Total Mean Change HR
Total Mean Change SRL	-.15

No significant association was found between total mean change SRL and total mean change HR.

6. DISCUSSION

Firstly, the findings of the study will be summarised and discussed. Limitations of the study will then be presented in relation to the interpretations as well as addressing the associated theoretical and clinical implications of the study. Strengths of the study will be presented before future directions for research are suggested.

6.1. Summary of Results

6.1.1. Hypothesis 1

As predicted in Hypothesis 1(a) both total cognitive empathy and affective empathy subscales of the VERA(ii) demonstrated high overall internal consistency (Field, 2005), providing evidence for the reliability of the measure and supporting previous findings by Young et al. (2008) and Terry et al. (2009). Hypothesis 1(b) was also supported with each vignette also displaying high overall internal consistency.

Additional exploration into the internal reliability of the total cognitive and affective empathy subscales of the VERA(ii) suggests that; for cognitive empathy, Interview 4 – Child Sexual Abuse, is not measuring cognitive empathy to the same degree as the rest of the interviews within the scale, and should this be removed the reliability of the scale would increase. The item-total correlations were above $r = .8$ for all of the items within the affective empathy subscale showing that there is a lot of repetition within the scale (i.e. each interview is not bringing anything additional to the overall scale). Subsequently demonstrating that there is a high level of redundancy of the items' contribution to the overall scale, supported by the high overall alpha ($r = .94$). Morgan and Griego (1998) report that a very high alpha, for example greater than $\alpha = .90$, suggests that items comprising the scale are repetitious, such that there are more items in the scale than are necessary for a reliable measure of the construct.

6.1.2. Hypothesis 2

Hypothesis 2 which predicted that total cognitive and affective empathy scores would be significantly positively correlated with an existing global empathy measure; the IVE-Empathy self-report measure, was not supported. Associations between both total cognitive and affective empathy constructs and the IVE-Empathy were demonstrated in the predicted direction (i.e. as higher scores on the VERA(ii) subscales increased, scores on the IVE-Empathy also increased) although neither correlation was significant. This finding is consistent with that found by Terry et al. (2009) who administered the original VERA and IVE-Empathy to an adult male community sample, for cognitive empathy but not affective empathy. However, the relationship between affective empathy and IVE-Empathy in the current study was only marginally non-significant ($p=0.057$). Nevertheless, as no significant correlations were found in the current study at the $p < 0.05$ level this brings into question the construct validity of the VERA(ii); (i.e. the relationship

between the VERA(ii) and other variables predicted to be associated with empathy). As is suggested by Terry et al. (2009) however, the failure to find a significant association with an existing global empathy measure may not indicate a lack of evidence for the validity of the VERA(ii), as other studies have also failed to demonstrate an association between victim empathy and global empathy measures (e.g. Tierney & McCabe, 2001).

6.1.3. Hypothesis 3

As argued in Chapter 1, in order to overcome the limitations of using self-report measures of empathy in isolation, physiological measurements (i.e. SRL and HR) were also included in the study. Predictions as stated in Hypothesis 3, that total cognitive and affective empathy would be negatively correlated with SRL and positively correlated with HR during stimulus presentation, were not supported. Initial findings of SRL and HR indicated that overall a decrease in SRL and an increase in HR occurred from baseline to participants being presented with the victim interviews, as has been demonstrated by Anastassiou-Hadjicharalambous and Warden (2007) and Fernández et al. (2012). This indicates that in general participants were objectively experiencing physiological arousal when viewing victims talking about their experiences compared to a resting period. However, additional analysis found this change in arousal was only significant for SRL but not HR. The difference in SRL was calculated to show a medium effect size, indicating that overall the vignettes are eliciting a tangible reaction in SRL. Additional correlations between mean total SRL and mean total HR, and total mean change in SRL and total mean change in HR were conducted to assess for any additional relationship between the physiological measures. Findings suggested a negative association between the two measures, as one would predict, although these were not significant. With regards to cognitive and effective empathy these findings indicate that although participants are displaying psychological arousal in reaction to viewing the vignettes, these were not correlated with their level of change in physiological arousal from baseline level.

6.1.4. Hypothesis 4

There was no significant negative association with EPQ-R Psychotism and either total cognitive or affective empathy as predicted in Hypothesis 4(a), suggesting that anti-social personality traits are not related to cognitive and affective empathy in this sample. A significant negative correlation was found between total cognitive empathy and violent attitudes with a small effect size (Cohen, 1988), but not affective empathy, partially supporting hypothesis 4(b). This indicates that the less one endorses violence, in other words the lower one's violent attitudes (i.e., cognitive empathy), the greater one's ability to understand the victim's distress and predict how she may feel. However, there is no association between violent attitudes and one's ability to experience an appropriate emotional response to viewing a victim being interviewed (i.e., affective empathy). Hypothesis 4(c) which stated that there would be a significant positive correlation between total cognitive and affective empathy and social comprehension / theory of mind was also partially supported, with a significant association being reported for cognitive

empathy with medium effect (Cohen, 1988), but not affective empathy. This suggests that as one's ability to interpret social interactions increases, one's ability to imagine a victim's distress and predict how she may feel also increases. Hypothesis 4(d) which stated that there would be a significant positive correlation between total cognitive and affective empathy and neuroticism was not supported. Finally, no significant positive correlation was found between either total cognitive or affective empathy and estimated full-scale IQ failing to provide support for Hypothesis 4(e), suggesting that intelligence is not associated with one's ability to empathise, either cognitively or affectively. However, it is of note that the distribution of the sample for intelligence was skewed, with the average estimated full-scale IQ for the sample indicating high average intelligence.

6.2. Interpretation of Results

This section will discuss the findings before addressing these in light of the limitations of the study and the associated theoretical and clinical implications.

Total scores for both cognitive and affective empathy demonstrate that participants are subjectively reporting that they are experiencing an empathic response in relation to the watching the victims being interviewed about their experiences, over a range of non-criminal and criminal offences. The distribution of total scores for both constructs ranged greatly with some participants scores showing that they experienced no cognitive empathy at all for the victim and almost no affective empathy, yet for other participants their scores suggested that they were indeed highly empathic, both at a cognitive and an affective level. This finding will be discussed in greater detail in the limitations section. Overall, total mean scores for the sample were greater for cognitive empathy compared to affective empathy, with further analysis of the difference between the constructs confirming that this was significant demonstrating a large effect size, consistent with findings presented by Young et al. (2008). This suggests that participants understood the victims' distress, were able to imagine the victims' perspectives and consequently predict how the victim may feel to a greater extent than they experienced an appropriate change in emotion in response to the victim.

The current study aimed to validate the VERA(ii) using both subjective self-report measures and objective physiological measures of arousal, combining the methodologies to overcome the limitations of using each in isolation. However, although data were obtained for the sample, demonstrating that participants were experiencing distress associated to viewing the VERA(ii), by the self-report and physiological methods independently, no convergence between the two was found. Upon reflection this finding is not unexpected as although employing the use of the methods in combination has been increasingly advocated within empathy research (Eisenberg & Fabes, 1990; Baron-Cohen, 2011), to date findings of convergence between these types of measures has been mixed (Fernández et al., 2012).

One possible explanation to account for the discordance between self-reported distress and physiological levels of arousal, as has been suggested in previous research, focuses upon the low levels of emotional arousal generated (Hodgson & Rachman, 1974; Gudjonsson, 1980; Fernández et al., 2012). For example, although physiological arousal is occurring this may not be occurring at a level sufficient enough to be associated with the emotions stated through self-reports, supporting Hodgson and Rachman (1974) argument that the concordance between the two methods is affected by the intensity of the emotional arousal. With regards to the current study therefore, although SRL and HR demonstrated that physiological arousal was occurring in response to viewing the vignettes, this may not have been produced at a level great enough to be considered a tangible empathic response. That the current study found a significant difference in SRL but not HR is again not unexpected as EDA has been stated to be one of the most sensitive physiological measures available (Gudjonsson, 1980) as a general arousal indicator whereas HR is associated with more subtle physiological states of interest to the researcher (Kreibig, 2010). SRL displayed an increase from resting to viewing the stimulus, demonstrating a significant medium effect size, indicating that the vignettes were eliciting a reaction compared to a resting stimulus free period. However, as this failed to converge with the VERA(ii) cognitive and affective empathy subscales this reaction cannot be considered to be either a tangible cognitive or an affective empathic reaction. If the arousal cannot be considered to be an empathic reaction, what then may this be?

As Fernández et al. (2012) concluded in their study, viewing films that tap into emotions with higher subjective arousal will result in the individual experiencing increased physiological activation. They reported that emotions such as fear and anger have been shown to elicit greater increases in physiological arousal than emotions such as tenderness and sadness. Furthermore they found a significant convergence between subjective self-reports of fear and anger and physiological reactions although this was not so for sadness, despite sadness demonstrating an increase in physiological arousal. They state that this may be so as fear and anger create greater sympathetic nervous system activation and arousal to prepare the individual for a more active behavioural response (i.e. fight or flight). Conversely, they argue that emotions such as tenderness and sadness often have the function of communicating and influencing others and so would not require such intensive physiological responses to be generated. This is an important finding with regards to empathy and the current research as empathy is a complex emotional construct. Reacting empathically to another may potentially activate a range of emotions, which may depend on the individual with whom one is empathising. It may therefore be that for participants in the current study the VERA(ii) was tapping into emotions such as sadness for, and tenderness towards, the victim as an empathic response as subjectively reported in the experimental setting, and which in real life would be communicated back to the victim. However, as these emotions do not cause intense physiological responses to be generated they subsequently do not produce a convergence with the cognitive and affective empathy that is self-reported. Therefore although a convergence was not demonstrated between the physiological and self-report measures in the current study,

it is argued that this does not necessarily indicate that the individual has not responded empathically towards the victim.

A similar explanation to that presented above (i.e. that the level of physiological arousal one experiences may be dependent on what combination of emotions are generated as part of the empathic process), includes the individual's previous experiences; the context which the other is describing his/her distress, for example in a real life or experimental situation; the type of offence/incident being recalled; and characteristics of the victim, as well as current stresses affecting the observer (Baron-Cohen, 2011; Marshall et al., 1995). These are all factors that may influence how one interprets another's distress. For example, if a vignette is personally salient to an individual due to the individual being the victim of a similar incident, having witnessed a similar incident, or knowing someone who has also experienced a similar incident, then one would expect this to generate a greater objective automatic physiological reaction as well as a greater subjective empathic response. Additionally, one would expect to observe a greater physiological and self-reported empathic response to viewing a real life incident in comparison to being shown video clips of the same situation within an experimental setting. Focusing upon these examples it is suggested that the individual's increased physiological arousal and subsequent empathic response may be created by the activation of a different range of emotions for example, fear and / or anger, in addition to sadness and tenderness. As noted previously these emotions would indeed generate a more intense physiological reaction than when correlated with subjective self-reports of empathy and may produce a convergence between the two methods. On the contrary as Vlahou et al. (2011) argue, stimuli that are salient to the participant may cause the participant to engage in emotional reappraisal as a means to regulate their emotional response, which would also create the discrepancy as found in the current study. Conversely, if an individual struggles to identify with the victim for example, if the individual has difficulty connecting to the victim due to gender and age differences, as suggested by Terry et al. (2009), this may limit the individual's ability to experience the associated emotional reaction in response to how the victim is feeling. In this situation one would expect to observe lower levels of physiological arousal. This may have been a factor in the current study which employed non-offending male participants aged between 20 and 60 years old to watch videos of young adult female victims being interviewed about their experiences in an experimental setting. However, the individual may nevertheless report how they think they should feel in response to the situation, which would then also create discordance between the two methods.

As has been argued previously, participants may have rated their subjective self-reports of empathy as greater than they actually felt due to a want to appear socially desirable (Young et al., 2008; Grady et al., 2011). However, this argument is not supported for the population of the current study as no significant associations were found between socially desirable responding (as measured by the EPQ-R Lie subscale) and the VERA(ii) total cognitive and affective empathy subscales. The discordance between the measures, particularly the lack of an intense physiological arousal to the VERA(ii), may instead possibly be explained by society's

increased tolerance to viewing distressing stimuli. It is argued that in general, society is increasingly being exposed to more violent images, portrayals of crimes and victim experiences on the T.V., and in films. Consequently, it may be that people are developing a tolerance to viewing and hearing about distressing experiences, therefore when presented with interviews of victims talking about their experiences in the VERA(ii), this stimuli is failing to activate an appropriate physiological response in relation to participants self-reported distress, as may once have been the case when exposure to such things was infrequent. Or, as Baron-Cohen (2011) suggests people may simply be over-estimating how empathic they are when subjectively reporting this. This again suggests that possibly participants are having an empathic response to the victim, as highlighted in their self-reports, yet this is not being demonstrated in the physiological measurements. An alternative explanation to address the discordance between measures may be explained with regards to cognitive processes (i.e., memory and attention).

The findings of the current study are similar to those reported by Anastassiou-Hadjicharalambous and Warden (2007) who found that when presenting evocative film clips to children, physiological reactions (HR) increased significantly relative to a resting baseline period, but convergence between verbal self-report of the emotional reaction experienced in response to the film clip and change in HR only occurred at chance levels. Anastassiou-Hadjicharalambous and Warden (2007) argue that the reasons for this result may have been methodological. They state that the time delay from recording the objective physiological reaction to the subjective self-report may have allowed cognitive processes (i.e. memory and attention) to interfere with how one interpreted the stimuli, subsequently affecting one's subjective reports. In other words, as a result of the delay from viewing the stimuli to reporting one's emotional reaction at the time of viewing, subjects are therefore required to recall or reconstruct their emotional experience, which may be different to that which was experienced at the time, resulting in the discordance between physiological and self-report measures. With regards to the current study it is possible that such cognitive processes played a part in the difference between self-reported empathy and physiological reactions also. However, although both studies administered self-report measures immediately after viewing the stimulus, the length of time of the stimulus were notably different; the evocative film clips presented by Anastassiou-Hadjicharalambous and Warden being 22 minutes in duration compared to approximately 3 ½ minutes in the current study. Despite this however, cognitive processing of the stimuli by participants may also account for additional findings in the current study.

Significant associations were demonstrated between total cognitive empathy and violent attitudes with a small effect size, as well as total cognitive empathy and social comprehension with medium effect size. In turn these findings suggest that as participant's attitudes to violence decrease, their ability to cognitively empathise with the victim increases (i.e. the lower one's violent attitudes, the greater one is able to imagine the victim's perspective and accurately predict how she may feel as a result). Additionally, it has been demonstrated that as the participant's level of social comprehension increases so too does his ability to cognitively empathise with the victim (i.e. the greater one is able to understand and interpret social

situations, the greater one is able to understand how the victim may be feeling). Together these findings suggest that overall the sample holds beliefs that do not justify the use of violence and they are also able to understand how one should behave in social situations. In combination therefore this assists participants to imagine how distressing the incident may have been for the victim thus they are able to accurately identify how the victim may have been feeling. However, although the participant is able to understand how the victim may be feeling at the cognitive level as a result of these factors, they may nevertheless fail to personally identify with the victim such that an affective response to her distress would not be generated. If this is so, a strong physiological reaction would not be generated that would correspond to subjective self-reports.

Overall, it has been suggested that validation of the VERA(ii), using a dual method approach was unsuccessful, indicating that the VERA(ii) is not successfully tapping into victim empathy as it purports, based on the discordance between the two methods. However, explanations have been suggested in an attempt to explain why this discordance may have occurred. Possible explanations include: the complexity and range of emotions associated with the empathic response; the context in which one is placing the stimuli, thus how one is interpreting the victim's distress and subsequently reacting to this; and cognitive processing of the stimuli. Furthermore, it has been argued, as Terry et al. (2009) suggested, that failure to find an association between an existing global empathy measure and the VERA(ii) may be a consequence of the scales measuring different aspects of empathy; the IVE-Empathy measuring global empathy while the VERA(ii) measures victim empathy. A review of the limitations of the study with regards to the current findings will now be discussed.

6.3. Limitations, Theoretical and Clinical Implications

The main limitation identified within the current study concerns the somewhat artificial nature of the vignettes. The main objective in the development of the VERA(ii) was to enhance the 'realism' of the interviews, which was deemed to be lacking in the original audio VERA. Subsequently video vignettes utilising professional actresses were produced. However, reports from participants (after being debriefed, and by some while viewing the vignettes) typically stated that they felt that the interviews seemed "fake", reporting that "the victims are clearly actresses" as "real victims would not act in that way". As such, they struggled to empathise with the 'victim' to the same degree that they might had they believed that the victim was real. One participant felt so strongly about this that he reported, upon being debriefed, that such was his disbelief that the victims were real and recalling real personal experiences, that he was not able to report that the victim experienced any distress, resulting in a cognitive empathy score recorded as 0. Baron-Cohen (2011) would explain this reaction as the individual experiencing a 'temporary empathy deficit'. This would not be considered a permanent deficit as the current study found no correlation between EPQ-R Psychopathy and total cognitive and affective empathy to suggest this. To explain, as a result of the artificial nature of the VERA(ii), Baron-Cohen would argue that the empathy circuit of the participant may have been temporarily shut down such he was unable to empathise with the victim. This can also be applied to how one

perceives the context of the vignettes, as discussed previously. This has consequences such that if the empathy circuit in the brain is shut down the sympathetic nervous system would not be activated and one would not have a resulting physiological reaction. If this was occurring at different levels across the population sample this would then produce less overall intense physiological reactions. However, a range of total cognitive and affective empathy scores were reported providing evidence that not all participants felt this way. However, this may have been a factor in the discrepancy between the physiological reactions and self-reported empathy scores. Additionally, this may also account for the difference between total cognitive and affective empathy scores.

To clarify, if a participant is questioning the 'realism' of the vignettes whilst the stimulus is being displayed, it is argued, as is suggested above, that he will be engaging in cognitive processing that will be taking cognitive capacity away from attending to the victim; specifically understanding and processing the content of what she is saying, as well as her body language and facial expressions which provide evidence of her distress. The participant may instead be engaged in looking for cues for example, how the victim acts in the interview and in the setting of the interview, to confirm their assumption that the victim is an actress. If the participant is engaging in cognitive processing of this sort then one would expect a low physiological reaction in response to the vignettes. Where physiological reactions are occurring these may be related to doubt and curiosity as opposed to distress for the victim. However, the participants may subjectively self-report that they felt empathy towards the victim had she been real. This has two consequences for the current study. Firstly, if the participant has shifted his attentional resources to assist in identifying if the victim is an actress then he will have generated a low physiological response, one that does not underlie an empathic response. However, when presented with the cognitive and affective subscales he may have chosen to respond to these how he predicts he would feel, having not been told by the researcher that the victim was real or not. This would produce a discordance between his physiological measures and his self-reports, as was found in the current study. Secondly, this may also explain the difference between total cognitive and total affective empathy for the current sample. For example, if the participant is subjectively reporting how he perceives the victim would be feeling if she were real then he is likely to score high on cognitive empathy. However, in contrast he may respond to the affective empathy subscale with how he felt when watching the vignette when he believed that the victim was an actress, which would elicit a lesser response than if he were describing how he may have felt had he believed that the victim was real.

Theoretically this can be explained using Marshall and Marshall's model (2011). For example, the participant was able to achieve the first two stages, emotion recognition and perspective taking, as he was able to imagine what emotion would be felt by a real victim, assisted by the content of the vignette, thus reporting this in the cognitive empathy subscale. However, he fails to achieve the latter stages as he is unable to replicate the emotion of the 'victim' due to his belief that she is not truly experiencing this emotion herself. In other words, how can one replicate the emotion of another if he does not believe the other is really experiencing this? This

would then account for the lower scores on the affective empathy scale, as well as a less intense physiological reaction as discussed above.

The second limitation of the study concerns the order in which the vignettes were presented. The vignettes were arranged in an order as constructed by the manufacturer commissioned to produce VERA(ii), such that they were displayed in the order; arson, car accident, assault, child sexual abuse and rape. This is not consistent with previous studies and therefore may present as a confounding factor in any comparisons made with the original VERA. The order of VERA(ii) therefore appears to increase in severity beginning with the non-criminal vignettes and ending with the criminal vignettes. Spontaneous comments from some participants during administration of the VERA(ii) highlighted that towards the end of the series they were accurately able to predict the subject of the latter vignettes (i.e. that the last vignette was an interview with a victim of rape). Consequently this methodological limitation may partly account for the discrepant findings in the current study. As the vignettes were increasing in severity this may have allowed the participants to prepare themselves for what was to come, they may have been able to implement cognitive strategies to protect themselves from the resulting distress that they felt the latter vignettes may cause, so regulating their emotional response. This would then account for the discordance between their physiological reactions and subjective self-reports, as argued by Vlahou et al. (2011).

Alternatively, participants may have become habituated to the vignettes over time, despite their increasing severity. Participants may have been socialised to the structure of the interviews over the number of presentations such that they were conditioned as to what to expect, resulting in a decrease in physiological reaction over time, as is described by the orienting response (Graham, 1979; Dawson et al., 2000). As participants predicted the vignettes were increasing in severity their reactions may have become less responsive to the material presented, which may not have occurred if the order of the vignettes were presented in a more random fashion, making it unlikely that they could predict what they were about to view next. This may have had a result on the severity of their subjective self-reports, particularly with regards to affective empathy, as well as physiological reactions if they had they been presented with a criminal vignette first for instance.

An additional limitation of the study is attributed to the focus of the analysis. The study focused upon analysing total cognitive and affective empathy and therefore used total mean and total change in physiological measures from which to investigate any associations with subjective self-reports. However, as has been reported in previous studies with the original VERA, some vignettes have been shown to evoke a greater response than others (i.e. the criminal rape vignette in comparison to the non-criminal car accident; Young et al., 2008). Combining the physiological measures across vignettes may therefore have lost some of the detail of the reactions for the vignettes that have been subjectively reported in previous studies to be more distressing. Using mean physiological measurements may therefore have contributed to the

lower arousal levels recorded which may in turn have affected the concordance between this and the subjective measurements in the current study.

Although the current study aimed to overcome the limitations associated with self-report measures of empathy by combining physiological measures, it nevertheless incorporated an existing global measure of empathy as a means to assess for construct validity of the VERA(ii). This failed to provide construct validity against the VERA(ii). However, it is suggested that this failure is a consequence of a global measure of empathy attempting to validate a victim specific measure, as argued by Terry et al. (2009). Due to this it is suggested that the current study, although aware of the limitations regarding the validity and reliability of existing victim empathy measures should have nevertheless incorporated a victim specific empathy measure in addition to the IVE-Empathy, as a means to assess for construct validity of the VERA(ii).

Reliability analysis found both total cognitive and affective empathy subscales to have high internal consistency. However, further analysis into the composition of the items within each suggests that for cognitive empathy removal of Interview 4 – Child Sexual Abuse would increase the reliability of the subscale. For affective empathy further exploration into the composition of the scale is needed as some items are redundant. It is suggested therefore that further exploration of the reliability of the subscales is conducted with regards to the continued development of the VERA(ii) and assessing its validity.

Finally, although the current sample was calculated to be large enough to detect associations between the identified variables with a power of 0.8 and a probability level of $p < 0.05$, nevertheless it is argued that employing a larger sample would be beneficial. For example, had a larger sample been utilised the distribution of the data for all scales may have been normal thus allowing parametric tests to be applied in the analysis. This would have increased the sensitivity of the analysis and power of the correlations conducted opposed to the non-parametric tests used where this is reduced due to ranking of the data. Furthermore, this would have overcome the limitation of non-parametric tests as stated by Field (2005), that the data may not generalise beyond the study sample.

6.4. Strengths

Despite the limitations outlined, the study also had a number of strengths. The main strength of the study was incorporating a physiological measure to assist with the validity of the measure, attempting to overcome limitations of self-report measures highlighted by Terry et al. (2009), following validation of the original audio VERA. Additionally, VERA(ii) attempted to increase the authenticity of the previous measure by introducing a visual aspect to the victim report in an attempt to increase the ecological validity of the tool in comparison to the original audio version of the VERA, further building on the usefulness of the written vignette alone. Applying a dual-method approach was successful in establishing that participants are indeed having an

emotional response to the vignettes as reported. However, it has highlighted that this response is not tapping in to cognitive or affective empathy as predicted.

6.5. Future Research

Recommendations for future research in light of the current findings and the limitations of the study follow.

Overall, the current study has provided support to the argument that cognitive and affective empathy are distinct concepts, therefore it is necessary for future victim empathy research to measure these independently. Clinically, it is suggested that treatment programmes may therefore wish to focus to a greater extent on developing affective empathy skills of offenders. For example, increasing one's ability to experience the distress of the victim, in addition to being able to understand the victims distress, may be a greater deterrent for offenders. Furthermore, as the current failed to find a convergence between physiological measures and self-report measures, it is suggested that future studies may wish to employ more sensitive physiological measures. This could be attained by utilising multiple cardiovascular and electrodermal measures, as well as introducing respiratory measures for example. As suggested was the case by Fernández et al. (2012) in their emotion research, those utilised in the current study may not be sufficient to represent complex physiological reactions in relation to the construct being measured.

In response to participant feedback, to increase the ecological validity of the VERA(ii), and overcome the artificial feel of the vignettes, it is recommended that future studies utilise real victims. However, due to the associated ethical implications, whether this is truly feasible is questionable. With this in mind it is therefore suggested that if the VERA(ii) was to be utilised in future studies then participants would be informed prior to the administration that the victims are indeed actresses. It is hypothesised that being more directive, thus instructing participants to focus on the content of what the 'victim' is saying, may assist the participant to focus his attention more fully on the content of the victim interview. This will potentially not only increase his attention but will also assist with his cognitive processing to the content of the interview, as opposed to his attention and cognitive processing being engaged in assessing if the victim is real. This may have the resulting physiological response associated with a subjective empathic response, thus a significant association may be found between physiological and self-report measures as was hypothesised in the current study.

To address the possibility that the context of the vignettes may be influencing participants reactions, future research may wish to incorporate ways to identify if participants found the vignettes to be particularly salient, or not. This could be achieved by incorporating additional self-report measures that assess for traumatic experiences for example. Alternatively this information could be gathered by qualitative methods such as utilising a semi-structured interview administered after the VERA(ii) has been completed. Using the latter method, one

would also be able to ascertain if participants had employed any strategies, cognitive or otherwise, as a means to regulate their emotional reaction to the vignette.

With regards to future research overcoming the limitation associated with the order that the vignettes were displayed, amending this from a fixed order to a random order would successfully rectify this. However, employing a random order to the presentation of the vignettes would create a large number of permutations therefore the sample size of the study would need to increase dramatically for this to be applied successfully. Increasing the sample size of the study would also be advantageous as this would also increase the power and sensitivity of the study to detect significant effects.

Future studies may wish to widen the focus of the analysis to assess for change in physiological reactions across each vignette individually as well as cumulatively, as was the case in the current study. This would allow for more detail to be collected to assess if, as has been shown by Young et al. (2008) and Terry et al. (2009), some vignettes elicit stronger reactions than others. Additional more sensitive analytic methods may also be incorporated. For example, when examining the physiological data one may wish to look at the number of peaks in SRL indicating arousal to ascertain if these are related to specific emotive words or phrases within each vignette. This may give an indication of what emotions are being generated and recorded by the physiological measurements, which will help to ascertain what emotions are being generated in relation to the VERA(ii).

In light of the additional reliability analysis of the VERA(ii), reliability of both total cognitive and affective empathy subscales should be further addressed to assist in the development of the scale (i.e. where items are redundant these should be removed as a means to refine and improve the scale). Improving the reliability of the scale will subsequently assist with assessing its validity.

As construct validity failed to be achieved with an existing global empathy measure, future research may also wish to include a victim specific measure for example, the Victim Empathy Distortions Scale (Beckett and Fisher, 1994) which has demonstrated good levels of internal reliability, test-retest reliability and discriminant validity to date.

Due to the failure of the current study to find an association between physiological and self-report measures of empathy and the mixed results reported from convergence studies using these methods, both in relation to empathy and emotion research, maybe future research should look towards incorporating alternative objective measures of arousal. For example, as has been employed by Baron-Cohen and colleagues in their research in identifying the 'empathy circuit', maybe future research could employ the use of brain imaging studies in the objective analysis of empathic responding to stimuli such as the VERA(ii) in comparison to self-report measures. Although this would be a costly endeavour in the continued validation of the

VERA(ii) the results would have far reaching theoretical and clinical implications with regards to victim empathy research and the developing treatment programmes for offenders.

The current study utilised a non-offending adult male community sample to validate the VERA(ii). Although evidence to support the validation of the study was found to be limited, nevertheless this should not prevent discriminant validity of the measure being sought with an offending population. Findings may differ significantly with this group due to a number of factors, including how they interpret the vignettes due to context, situation and previous experiences, as stated previously for this sample. For example, research has shown that a high proportion of sex offenders have suffered child abuse in comparison to the normal population. This may then effect their responses, physiological or subjective, with regards to the child sexual abuse vignette as they may identify through a shared experience with the victim. Conversely, the sex offender may not identify with the rape victim. It is therefore important to continue to validate VERA(ii) with this population for which it is designed to be used therapeutically. Additionally, future research should also use a female population as research has shown females to display greater levels of empathy than males which may have an effect on physiological responding in relation to subjective self-reports of empathy. Using a female sample may also address some of the limitations with regards to the context of the vignettes. Finally, although the current study found no association between intelligence and empathy, intelligence for the sample was skewed towards high intelligence indicating that the sample was somewhat selective. Consequently, future studies should aim to recruit a more representative sample to more accurately determine if this does have an effect on the main hypotheses.

6.6. Conclusions

The current study was concerned with the validation of a new victim empathy measure, the Victim Empathy Response Assessment (ii) (VERA(ii)). Analysis of the scale found the VERA(ii) total cognitive and affective empathy subscales to have good overall internal consistency, providing evidence of test-retest reliability with previous studies (Young et al., 2008; Terry et al., 2009). However, further reliability analysis suggests that the subscales may be improved. Evidence was also found supporting the previous finding that cognitive empathy is greater than affective empathy with regards to VERA(ii). Construct validity assessed against an existing global empathy measure was not supported. Convergent validity was assessed by correlating physiological measures of arousal, recorded when viewing the VERA(ii), with VERA(ii) total cognitive and affective subjective self-report scores. However, no significant results were found. No significant associations were found between total cognitive or affective empathy and personality variables or intelligence. Significant associations were found between total cognitive empathy and violent attitudes and social comprehension to suggesting that the less one endorses violence, and the greater one is able to accurately interpret social interactions, the greater one is able to understand a victim's distress and predict how she may feel. However, there was no association between either violent attitudes or social comprehension and affective empathy, suggesting that although one may be able to understand a victim's distress and

predict how she may feel, one does not have an appropriate emotional response to the victim's distress. Overall, the results of the current study have provided limited evidence for the validity of the VERA(ii). Numerous limitations of the study have been highlighted which may have affected the current findings that have been discussed with regards to theoretical and clinical implications. Recommendations for future research have subsequently been presented.

In summary, it is concluded that victim empathy is a complex construct that is difficult to measure through both objective and subjective methods, with the current study demonstrating that the VERA(ii) is perhaps not tapping sufficiently into the empathy construct as it purports. However, this may be so due to limitations of the current study. Consequently, further research to develop the reliability of the VERA(ii) and validate the measure is needed, using an offending sample and possibly incorporating additional physiological measures of arousal.

7. REFERENCES

- Alexio, P.A., & Norris, C.E. (2000). Personality and moral reasoning in young offenders. *Personality and Individual Differences*, 28, 609-623
- Anastassiou-Hadjicharalambous, X., & Warden, D. (2007). Convergence between Physiological, Facial and Verbal Self-Report Measures of Affective Empathy in Children. *Infant and Child Development*, 16, 237-254.
- American Psychological Association (1986), as cited by D.W. Tierney & M.P. McCabe. (2001). *The Assessment of Denial, Cognitive Distortions, and Victim Empathy Among Pedophilic Sex Offenders: An Evaluation of the Utility of Self-Report Measures*. *Trauma, Violence & Abuse*, 2(3), 259-270.
- Baron-Cohen, S. (2011). *Zero Degrees of Empathy: A new theory of human cruelty*. London: Allen Lane.
- Batson, C., Fultz, J., & Schoenrade, P. (1987). Adults' emotional reactions to the distress of others. In N. Eisenburgh, & J. Stryer (eds.), *Empathy and its development* (pp. 163-184). Cambridge: Cambridge University Press.
- Beckett, R.C., & Fisher, D. (1994). *Assessing victim empathy: A new measure*. Paper presented at the 13th Conference of the American Association for Treatment of Sexual Abusers (ATSA), San Francisco, USA.
- Ben-Shakar, G., Bornstein, G., Hopfensitz, A., & van Winden, F. (2007). Reciprocity and emotions in bargaining using physiological and self-report measures. *Journal of Economic Psychology*, 28, 314-323.
- Blascovich, J., Vanman, E.J., Berry Mendes, W., & Dickerson, S. (2011). *Social Psychophysiology for Social and Personality Psychology*. (The SAGE Library of Methods in Social and Personality Psychology). (pp. 1-40). London: SAGE Publications Ltd.
- Boden, J.M., & Dale, K.L. (2001). Cognitive and affective consequences of repressive coping. *Current Psychology: Developmental, Learning, Personality, Social*, 20, 122-137.
- Bradley, M.M., & Lang, P.J. (1995). Measuring emotions: The self assessment manikin and the semantic differential. *Journal of Behaviour Therapy and Experimental Psychiatry*, 25, 49-59.

Briggs, D., Doyle, P., Gooch, T., & Kennington, R. (1998). *Assessing Men Who Sexually Abuse: A Practice Guide*. London: Jessica Kingsley.

British Psychological Society (2004). *A review of the current scientific status and fields of application of Polygraphic Deception Detection: Final report* (6 October 2004) from the BPS Working Party. Leicester: The British Psychological Society.

Brosschot, J.F., & Janssen, E. (1998). Continuous monitoring of affective autonomic response dissociation in repressors during negative emotional stimulation. *Personality and Individual Differences*, 25, 69-84.

Cacioppo, J.T., & Gardner, W.L. (1999). *Emotion*. *Annual Review of Psychology*, 50, 191-214.

Channon, S., & Crawford, S. (2000). The effects of anterior lesions on performance on a story comprehension test: left anterior impairment on a theory of mind-type task. *Neuropsychologica*, 38, 1006-1017.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. (2nd ed.). New York: Academic Press.

Cohen, J. (1992). *A power primer*. *Psychological Bulletin*, 112(1), 155-159.

Curwen, T. (2003). The importance of offence characteristics, victimisation history, hostility, and social desirability in assessing empathy of male adolescent sex offenders. *Sexual Abuse: Journal of Research and Treatment*, 15, 347-364.

Dawson, M.E., Schell, A.M., & Filion, D.L. (2000). *The Electrodermal System*. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (eds.), *Handbook of Psychophysiology*. (2nd ed.), (2007, p.159-181). Cambridge: Cambridge University Press.

de Vignemont, F., & Singer, S. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences* 10(10), 435-41.

Eisenberg, N., & Fabes, R.A. (1990). Empathy: Conceptualization, Measurement, and Relation to Prosocial Behavior. *Motivation and Emotion*, 14(2), 131-149.

Eysenck, H.J., & Eysenck, S.B.G. (1975). *Manual of the Eysenck personality questionnaire*. London: Hodder and Stoughton.

Eysenck, H.J., & Eysenck, S.B.G. (1991) *Manual of the Eysenck Personality Scales (EPS Adult)*. London: Hodder and Stoughton.

Eysenck, S.B.G., & Eysenck, H.J. (1978) Impulsiveness and Venturesomeness: Their Position in a Dimensional System of Personality Description. *Psychological Reports*, 43, 1247-1255.

Eysenck, S.B.G., & Eysenck, H.J. (1978). Impulsiveness and Venturesomeness in a detention centre population. *Psychological Reports*, 47, 1299-1306.

Eysenck, S.B.G., Eysenck, H.J. & Barrett, P. (1985) A revised version of the Psychoticism Scale. *Personality & Individual Differences*, 6 (1), 21-29

Eysenck, S.B.G., Pearson, P.R., Easting, G., & Allsopp, J.P. (1985) Age norms for Impulsiveness, Venturesomeness and Empathy in Adults. *Personality & Individual Differences*, 6 (5), 613-619

Farrington, D. (1998). Individual differences and offending. In M. Tonry (ed.), *The handbook of crime and punishment* (p. 241-268). New York: Oxford University Press.

Feldstein, J.C., & Gladstein, G.A. (1980). A comparison of the construct validity of four measures of empathy. *Measurement and Evaluation in Guidance*, 13, 49-57.

Fernández, C., Pascual, J., Soler, J., Elices, M., Portella, M.J., & Fernández-Abascal, E. (2012). Physiological Responses Induced by Emotion-Eliciting Films. *Applied Psychophysiology and Biofeedback*. Springerlink, Online First, 6th February 2012. Retrieved March 1, 2012 from: <http://www.springerlink.com>

Fernandez, Y.M., & Marshall, W.L. (2003). Victim empathy, social self esteem, and psychopathy in rapists. *Sexual Abuse: A Journal of Research and Treatment*, 15, 11-26.

Fernandez, Y.M., Marshall, W.L., Lightbody, S., & O'Sullivan, C. (1999). The child molester empathy measure: Description and examination of its reliability and validity. *Sexual Abuse: Journal of Research and Treatment*, 11(1), 17-32.

Feshback, N. (1975). Empathy in children: Some theoretical and empirical considerations. *Counselling Psychologist*, 5, 25-30.

Field, A. (2005). *Discovering Statistics Using SPSS*. (2nd ed.). London: Sage Publications Ltd.

Finset, A., Stensrud, T.L., Holt, E., Verheul, W., & Bensing, J. (2011). Electrodermal activity in response to empathic statements in clinical interviews with fibromyalgia patients. *Patient Education and Counseling*, 82, 355-360.

Fisher, D., & Howells, K. (1993). Social relationships in sexual offenders. *Sexual and Marital Therapy*, 8(2), 123-136. In S. Young, G. Gudjonsson, R. Terry, & J. Bramham. (2008). *Victim Empathy Response Assessment (VERA): The validation of a new measure for forensic inpatients*. *Journal of Forensic Psychology and Psychiatry*, 19 (2), 191-204.

Fisher, D., Beech, A., & Browne, K. (1999). Comparison of Sex Offenders to Nonoffenders on Selected Psychological Measures. *International Journal of Offender Therapy and Comparative Criminology*, 43(4), 473-491.

Freeman-Longo, R.E., Bird, S., Stevenson, W.F., & Fiske, J. (1995). *1994 nationwide survey of treatment programmes and models: Serving abuse reactive children, adolescent and adult sex offenders*. Brandon, VT: The Safer Society Programme and Press.

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, 7(3), 385-400

Gladstein, G.A. (1983). Understanding empathy: Integrating counseling, developmental and social psychology perspectives. *Journal of counseling psychology*, 30, 467-482. In T. Ward, R. Laws & S.M. Hudson (eds.). *Sexual deviance: issues and controversies* (p. 173). SAGE.

Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.

Grady, M.D., Broderson, M., & Abramson, J.M. (2011). The state of psychological measures for adult sexual offenders. *Aggression and Violent Behaviour*, 16, 227-240.

Graham, F.K. (1979). Distinguishing among orienting, defence and startle reflex. In H.D. Kimmel, E.H. van Olst., & J.H. Orlebeke. (eds.) *The orienting reflex in humans*. (p. 137-167). Hillsdale, NJ: Erlbaum.

Grubin, D. (2002). The potential use of polygraph in forensic psychiatry. *Criminal Behaviour and Mental Health*, 12, 45-55.

Grubin, D., & Madsen, L. (2006). Accuracy and utility of post-conviction polygraph testing of sex offenders. *British Journal of Psychiatry*, 188, 479-483.

Gudjonsson, G. (1980). Self-Reported Emotional Disturbance and its Relation to Electrodermal Reactivity, Defensiveness and Trait Anxiety. *Personality and Individual Differences*, 2, 47-52.

Hanson, R.K., & Scott, H. (1995). Assessing Perspective-Taking Among Sexual Offenders, Nonsexual Criminals and Nonoffenders. *Sexual Abuse: A Journal of Research and Treatment*, 7(4), 251-277.

Harpur, T.J., Hakstian, A.R., & Hare, R.D. (1988). Factor structure of the psychopathy checklist. *Journal of Consulting and Clinical Psychology*, 56, 741-747.

Hodgson, R., and Rackman, S. (1974). Desynchrony in measures of fear. *Behaviour Research and Therapy*, 12, 319-326.

Jolliffe, D. & Farrington, D. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behaviour*, 9, 441-476.

Kerem, E., Fishman, N., & Jorrelson, R. (1994). The experience of empathy in everyday relationships: Cognitive and affective elements. *Journal of social and personal relationships*, 18, 709-729.

Knopp, Freeman-Longo & Stevenson. (1992) as cited by A. Beech., D. Fisher., & R. Beckett. (1998). *Step 3: An Evaluation of The Prison Sex Offender Treatment Programme*. London: HMSO.

Kreibig, S.D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, 84(3), 394-421.

Lang, P.J.(1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*, 50(5), 372-385.

Langdon, R., Coltheart, M., & Ward, P. (2006). Empathic perspective-taking is impaired in schizophrenia: Evidence from a study of emotion attribution and theory of mind. *Cognitive Neuropsychiatry*, 11, 133-155.

Laws, D.R. (ed.) (1989). *Relapse prevention with sex offenders*. New York: Guilford.

Laws, D.R., Hudson, S.M., & Ward, T. (eds.) (2000). *Remaking relapse prevention with sex offenders*: London: Sage.

Lottridge, D.M. (2010). *Measuring Emotional Responses to Interaction: Evaluation of Sliders and Physiological Reactions*. PhD Thesis, University of Toronto. Retrieved December 28, 2011, from https://tspace.library.utoronto.ca/bitstream/1807/26292/1/Lottridge_Danielle_M_201011_PhD_thesis.pdf

Lykken, D.T. (1981). *A Tremor in the Blood: Uses and Abuses of the Lie Detector*. New York: McGraw Hill.

Mandryk, R.L., & Atkins, M.S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 65, 329-347.

Marcus, R., Roke, E., & Bruner, C. (1985). Verbal and non-verbal empathy and prediction of social behaviour of young children. *Perceptual and Motor Skills*, 60, 299-309.

Marshall, L.E., & Marshall, W.L. (2011). Empathy and antisocial behavior. *The Journal of Forensic Psychiatry & Psychology*, 22 (5), 742-759.

Marshall, W., Anderson, D. & Fernandez, Y. (1999). *Cognitive Behavioural Treatment of Sexual Offenders*. Chichester: Wiley.

Marshall, W., Hudson, S., Jones, R. & Fernandez, Y. (1995). Empathy in sex offenders. *Clinical Psychology Review*, 15(2), 99-113.

Marshall, W.L. (1999). Current status of North American assessment and treatment programmes for sexual offenders. *Journal of Interpersonal Violence*, 14, 221-239.

Marshall, W.L., & Barbaree, H.E. (1990). An integrated theory of the etiology and maintenance of sexual offending. In W.L. Marshall, D.R. Laws, & H.E. Barbaree (eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (p. 257-275). New York: Plenum.

Marshall, W.L., Hamilton, K., & Fernandez, Y. (2001). Empathy Deficits and Cognitive Distortions in Child Molesters. *Sexual Abuse: A Journal of Research and Treatment*, 13(2), 123-130.

Mauss, I.B., & Robinson, M.D. (2009). Measures of emotion : A review. *Cognition and Emotion*, 23(2), 209-237.

McGrath, M., Cann, S., & Konopasky, R. (1998). New measures of defensiveness, empathy and cognitive distortions for sexual offenders against children. *Sexual Abuse: A Journal of Research and Treatment*, 10, 25-36.

Moore, B. (1990). The origins and development of empathy. *Motivation and Emotion*, 14, 75-80.

Morgan, G.A. & Griego, O.V. (1998). *Easy Use and Interpretation of SPSS for Widows: Answering Research Questions with Statistics*. Lawrence Erlbaum Associates Inc.

Mulloy, R., Smiley, W.C., & Mawson, D.L. (1999). The impact of empathy training on offender treatment. *Focus on Corrections Research*, 11, 15-18.

Murphy, W.D., & Barbaree, H.E. (1988). Assessment of sexual offenders by means of erectile response: Psychometric properties and decision making. In D.R. Laws, S.M. Hudson, & T. Ward. (eds.) *Remaking relapse prevention with sex offenders* (2000, p. 285-302), London: Sage.

Norman, G.R., and Streiner, D.L. (2008). *Biostatistics: The Bare Essentials* (3rd ed.). People's Medical Publishing House: B.C. Decker.

O'Connell, M.A. (2000). Assessment and Community Monitoring. In D.R. Laws., S.M. Hudson., & T. Ward. (eds.), *Remaking relapse prevention with sex offenders* (p 285-302). London: Sage.

Pithers, W.D. (1994). Process evaluation of a group therapy component designed to enhance sex offenders' empathy for sexual abuse survivors. *Behaviour Research and Therapy*, 32, 565-570.

Pithers, W.D., & Laws, D.R. (1989). The penile plethysmograph: Uses and abuses in assessment and treatment of sexual aggressors. In D.R. Laws, S.M. Hudson, & T. Ward. (eds.) *Remaking relapse prevention with sex offenders* (2000, p. 285-302), London: Sage.

Regher, C., & Glancy, G. (2001). Empathy and its influence on sexual misconduct. *Trauma, Violence & Abuse*, 2(2), 142-154.

Rosenthal, M.Z., Gratz, K.L., Kosson, D.S., Cheavens, J.S., Lejuez, C.W., & Lynch, T.R. (2008). Borderline personality disorder and emotional responding: A review of the research literature. *Clinical Psychology Review*, 28, 75-91.

Ryan, G. (1999). Treatment of sexually abusive youth: the evolving consensus. *Journal of Interpersonal Violence*, 14, 422-436.

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films : A new tool for emotion researchers. *Cognition and Emotion*, 24(7), 1153-1172.

Scully, D. (1988). Convicted rapists' perceptions of self and victim: Role taking and emotions. *Gender and Society*, 2, 200-213.

Serran, G. (2002). The measurement of empathy. In Y. Fernandez (ed.), *In their shoes* (p 16-36). Oklahoma City, OK: Wood N' Barnes.

Shi, Y., Nguyen, M.H., Blitz, P., French, B., Fisk, S., De la Torre, F., Smailagic, A., Siewiorek, D.P., al' Absi, M., Ertin, E., Kamarck, T., & Kumar, S. (2011). *Personalised Stress Detection from Physiological Measurements*. Retrieved December 28, 2011, from http://www.ca.cs.cmu.edu/projects/stress_detect/stress_detect.pdf

Sparks, G.G., Pellechia, M., & Irvine, C. (1999). The repressive coping style and fright reactions to mass media. *Communication Research*, 26, 176-192.

Sweeney, K.T., Ward, T.E., & McLoone, S.F. (2012). Artifact Removal in Physiological Signals—Practices and Possibilities. *IEEE Transactions on Information Technology in Biomedicine*, 16(3), 488-500.

Taylor, S.E. (1991). Asymmetrical effects of positive and negative events: the mobilisation and minimisation hypothesis. *Psychological Bulletin*, 110, 67-85.

Terry, R., Gudjonsson, G., and Young, S. (2009). Psychometric properties and validity of the 'Victim Empathy Response Assessment' (VERA) in a community adult male sample. *The Journal of Forensic Psychiatry and Psychology*, 20(5), 752-768.

Thayer, R. (1970). Activation states as assessed by verbal report and four physiological variables. *Psychophysiology*, 7, 86-94.

Tierney, D.W., & McCabe, M.P. (2001). The Assessment of Denial, Cognitive Distortions, and Victim Empathy Among Pedophilic Sex Offenders: An Evaluation of the Utility of Self-Report Measures. *Trauma, Violence & Abuse*, 2(3), 259-270.

Tredoux, C., & Pooley, S. (2001). *Polygraph-Based Testing of Deception and Truthfulness: An Evaluation and Commentary*. Retrieved November 17, 2011, from <http://web.uct.ac.za/depts/psychology/plato/poly.pdf>

Varker, T., Devilly, G.J., Ward, T., & Beech, A.R. (2008). Empathy and adolescent sexual offenders: A review of the literature. *Aggression and Violent Behaviour*, 13, 251-260.

Verschuere, B., Crombez, G., De Clercq, A., & Koster, E.H.W. (2005). Psychopathic traits and autonomic responding to concealed information in a prison sample. *Psychophysiology*, 42, 239-245.

Verschuere, B., Crombez, G., De Clercq, A., & Koster, E.H.W. (2004). Autonomic and behavioral responding to concealed information: Differentiating orienting and defensive responses. *Psychophysiology*, 41, 461-466.

Vlahou, C.H., Vanman, E.J., & Morris, M.M. (2011). Emotional Reactions While Watching Graphic Medical Procedures: Vocational Differences in the Explicit Regulation of Emotions. *Journal of Applied Social Psychology*, 41 (11), 2768-2784.

Waid, W.M., Orne, M.T., & Wilson, S.K. (1979). Socialization, Awareness and the Electrodermal Response to Deception and Self-Disclosure. *Journal of Abnormal Psychology*, 88, 663-666.

Walker, J.S. (2005). The Maudsley Violence Questionnaire: initial validation and reliability. *Personality and Individual Differences*, 38, 187-201.

Walker, J.S., & Gudjonsson, G.H. (2006). The Maudsley Violence Questionnaire: Relationship to personality and self-reported offending. *Personality and Individual Differences*, 40, 795-806.

Ward, T., Keenan, T., & Hudson, S.M. (2000). Understanding cognitive, affective, and intimacy deficits in sexual offenders: A developmental perspective. *Aggression and Violent Behaviour*, 5(1), 41-62.

Warnock-Parkes, E., Gudjonsson, G.H., & Walker, J.S. (2008). The relationship between the Maudsley Violence Questionnaire and official recordings of violence in mentally disordered offenders. *Personality and Individual Differences*, 44, 833-841.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment

Weinberger, D.A., Schwartz, G. E., & Davidson, R.J. (1979). Low-anxious, high-anxious, and repressive coping styles: psychometric patterns and behavioural and physiological responses to stress. *Journal of Abnormal Psychology*, 88, 369-380.

West, P. (1982). *Reproducing naturally occurring stories: Vignettes in survey research*. Aberdeen, UK: Medical Research Council.

Wiehe, V. (1997). Approaching child abuse treatment from the perspective of empathy. *Child Abuse and Neglect*, 21, 1191-1204.

Young, S., Gudjonsson, G., Terry, R., & Bramham, J. (2008). Victim Empathy Response Assessment (VERA): The validation of a new measure for forensic inpatients. *Journal of Forensic Psychology and Psychiatry*, 19 (2), 191-204.

8. APPENDICES

APPENDIX 1



University of London

Institute of Psychiatry at the Maudsley

West London Mental Health 
NHS Trust

Many viewpoints. One vision.

INFORMATION SHEET FOR CONTROLS

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Victim Empathy Rating Assessment: The validation of a newly developed tool for assessing empathy among offenders.

We would like to invite you to participate in this original research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

The aim of the research

The main aim is to investigate the effectiveness of a newly developed measure of victim empathy. We are conducting the study with patients at Broadmoor Hospital and controls at the Institute of Psychiatry to see how well it measures victim empathy.

Why am I being asked to take part?

We are inviting control participants to take part who have registered their details with the MindSearch database for volunteers at the Institute of Psychiatry.

Do I have to take part?

No – it is up to you to decide whether or not to take part in the study. If you decide to you will be given this information sheet to keep and will be asked to sign a consent form. After this you are still free to withdraw from the research at any time without giving a reason and without consequence. A decision not to take part or to withdraw later on will not disadvantage you in any way. If you withdraw from the study you may request that we destroy all data that has been collected about you.

What will the study involve?

In total the study will take around 1.5 to 2 hours. During the study you will be asked to complete an assessment of empathy in which you will watch five people discuss their experiences of different incidents. After each video you will be asked to answer some questions about how you felt, and how you think they felt, about the incident. During this measure we would like to collect some physiological data too. This will involve you wearing several small electrodes which simply stick onto your skin. This measure will take about 30 minutes. We will also ask you to provide some basic background information and to complete a brief IQ test along with some other questionnaires taking about an hour in total.

Benefits of taking part

It is hoped that you will find it informative to complete these questionnaires about yourself, and by taking part in this research you will be helping us to understand more about victim empathy. Once the research measures are complete you will be receive £10 as a contribution for your time.

Risks of taking part

The electrodes are not intrusive and stick onto your skin. In some cases a small red mark may be left after the electrode is removed but this should not be long-lasting and will fade away. Some of the videos or questionnaire items may be sensitive, however this will be different for different people and, if necessary, you will be offered support and advice from the research team and your clinical team. A

Research Assistant will be available to answer questions and discuss the research with you. If you wish to stop your participation in the study at any point you may do so without giving a reason and without any consequence.

Confidentiality

All information collected will be anonymous and strictly confidential. Information you provide will not be unless it is relevant to your safety or the safety of someone else, in which case this will be reported to the research team and discussed further as appropriate.

The procedures for handling, processing, storing and destroying data are compliant with the Data Protection Act 1998. Data will be collected with only a participation number to identify it. Information linking participation numbers and patient names will be locked away and only researchers will have access to this for the purpose of collecting file data. Data will be stored securely for up to ten years.

Your participation is voluntary. If you change your mind, you are free to stop your participation and to have your data withdrawn without giving any reason.

What happens to my information?

This research data will be analysed by the research group at the Institute of Psychiatry. The results will be used to improve service provision and will assist with the development of better services. A report of the study findings can be sent to you once the research has been completed.

Results will be submitted to a journal for publication and will be discussed at professional conferences. Participants in the research will not be identifiable in any reports, journal articles or presentations. Further information on the research and results can be sought from the research team.

It is up to you to decide whether to take part or not. If you decide to take part you are still free to withdraw at any time and without giving a reason.

Who is conducting this research?

The research is being carried out by researchers from the Institute of Psychiatry, Kings College London.

Contact details

If this study has harmed you in any way you can contact Dr Susan Young at the Institute of Psychiatry on 020 7848 0002.

APPENDIX 2

NRES Committee London - Harrow

Level 7 Maternity, Room 019
Northwick Park Hospital
Watford Road
Harrow
HA1 3UJ

Telephone: 020 8869 3928

Facsimile: 020 8869 5222

19 April 2011

Dr Susan Young
Senior Lecturer in Forensic Clinical Psychology
Institute of Psychiatry
De Crespigny Park,
Denmark Hill
London
SE5 8AF

Dear Dr Young

Study title:	Victim Empathy Rating Assessment: The validation of a newly developed tool for assessing empathy among offenders.
REC reference:	11/LO/0314
Protocol number:	CSA/11/009

The Research Ethics Committee reviewed the above application at the meeting held on 12 April 2011. The committee were grateful to Ms. Helen Lister and Mr. Gareth Hopkins for attending to discuss the study.

Summary of the discussion

In discussion, the Committee noted the following ethical issues.

- More detail was required about the content of the video.
- Is there a potential for offenders to 'act up' to please the researchers?
- Perhaps participants needed to be made aware that they may be shown things which could be distressing?
- The questionnaire for controls should not include the wording 'the crime you committed'
- It was not clear why people over 60 would be excluded from the study

Ms. Lister and Mr. Hopkin were invited into the meeting room

1. The committee asked if the content of the video was the same as audio version that had been approved in a previous study. Ms. Lister said some changes had been made. She described the content of the video and confirmed that it is not particularly distressing as participants do not witness an actual event.
2. The committee asked if the controls would be asked the Gudjonsson questionnaire? Ms. Lister said no and confirmed that the participants would know what 'control' means as there had been no issue with this in previous

studies. The controls would be obtained from a list of individuals that had indicated they were willing to take part in research at the Institute of Psychiatry. The Gumtree website advert (included in the application) has been used successfully for other types of research by the Institute of Psychiatry.

3. The committee asked about the potential for the offenders to be eager to please? The researchers' said was a problem with existing measures but the content in the new measures should help reduce this potential.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below. In addition, the committee asked that the researchers spell check the Questionnaires.

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Questionnaire: Assault A and B	1	11 March 2011
Questionnaire: story comprehension test	1	11 March 2011
Advertisement	1	01 March 2011
REC application	1	01 March 2011
Participant Information Sheet: for controls	1	01 March 2011
CV - Professor G. Gudjonsson		01 March 2011
Participant Information Sheet: for patients	1	01 March 2011
Questionnaire: Car Accident A and B	1	11 March 2011
Questionnaire: MVQ	1	11 March 2011
Questionnaire: Rape A and B	1	11 March 2011
Questionnaire: Arson A and B	1	11 March 2011
Questionnaire: Child Sexual Abuse A and B	1	11 March 2011
Questionnaire: Gudjonsson B.A. Inventory	1	11 March 2011
Questionnaire: Adult Eysenck EPQ-SV	1	11 March 2011
Questionnaire: Adult Eysenck IVE	1	11 March 2011
Protocol	1	01 March 2011
Evidence of insurance or indemnity		
Investigator CV		01 March 2011
Participant Consent Form: for patients	1	01 March 2011
Participant Consent Form: for controls	1	01 March 2011
Covering Letter	1	01 March 2011
CV - Helen Lister		01 March 2011

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document “After ethical review – guidance for researchers” gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- ☐ Notifying substantial amendments
- ☐ Adding new sites and investigators

- ☐ Progress and safety reports
- ☐ Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

11/LO/0314

Please quote this number on all correspondence

With the Committee's best wishes for the success of this project

Yours sincerely

Dr Jan Downer
Chair

Email: alison.okane@nwlh.nhs.uk

*Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments
"After ethical review – guidance for researchers"*

*Copy to: Mrs Jennifer Liebscher
[R&D office for NHS care organisation at lead site]*

NRES Committee London - Harrow

Attendance at Committee meeting on 12 April 2011

Committee Members:

<i>Name</i>	<i>Profession</i>	<i>Present</i>	<i>Notes</i>
Mr David Anderson-Ford	Senior Lecturer	Yes	
Mrs Veronika Bernstein	Translator	No	
Mr Andrew Caunce	Chief Pharmacist	Yes	
Mr Kevin Coughlan	Retired	Yes	
Dr Jan Downer	Consultant Anaesthetist	Yes	
Miss Shelly Glaister-Young	Barrister	No	
Dr Richard Kaczmariski	Consultant Haematologist	No	
Mr Peter Laszlo	Retired Financial Services Compliance Officer	Yes	
Dr Mary Leung	Clinical Psychologist	No	
Ms Ann Malkin	Consultant Psychologist	No	
Ms Fatima Nathoo	Administrator	No	
Mrs Alison O'Kane	Co-ordinator	Yes	
Miss Maggie Otter	Chief Pharmacist	Yes	
David Wells	Laboratory Manager	No	
Mrs Doreen West	Lecturer	Yes	
Jim Wood	Retired IT Consultant	Yes	

APPENDIX 3

NRES Committee London - Harrow

Level 7 Maternity, Room 019
Northwick Park Hospital
Watford Road
Harrow
HA1 3UJ

Tel: 020 8869 3928
Fax: 020 8869 5222

20 June 2011

Dr Susan Young
Senior Lecturer in Forensic Clinical Psychology
Institute of Psychiatry
Institute of Psychiatry
De Crespigny Park, Denmark Hill
London
SE5 8AF

Dear Dr Young

Study title: Victim Empathy Rating Assessment: The validation of a newly developed tool for assessing empathy among offenders.
REC reference: 11/LO/0314
Protocol number: CSA/11/009
Amendment number: 1.1
Amendment date: 23 May 2011

The above amendment was reviewed in June 2011 by the Sub-Committee in correspondence

Ethical opinion

none

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Protocol	1.1	23 May 2011
Notice of Substantial Amendment (non-CTIMPs)	1.1	23 May 2011

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

11/LO/0314: correspondence

Please quote this number on all
--

Yours sincerely

Dr. J. Downer

Chair

E-mail: alison.okane@nwlh.nhs.uk

Enclosures: List of names and professions of members who took part in the review

*Copy to: Mrs Jennifer Liebscher
 Ms Angela Williams, West London Mental Health Trust*

NRES Committee London - Harrow

Attendance at Sub-Committee of the REC meeting on 22 June 2011

Sub-Committee of the REC in correspondence

Dr. J. Downer (Chair)
Ms. A. Malkin

Consultant Anaesthetist
Clinical Psychologist

APPENDIX 4

How did you feel about the arson when watching the video?		Not at all	A little	Somewhat	A lot	Very much so
1	How Worried did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	How Sad did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	How Upset inside did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	How Distressed did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	How Thrilled did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	How Sorry (regret that the incident happened) did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	How Disturbed did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	How Angry do you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	How Interested did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	How Happy did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11	How Disgusted did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12	How Excited did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13	How Frightened did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

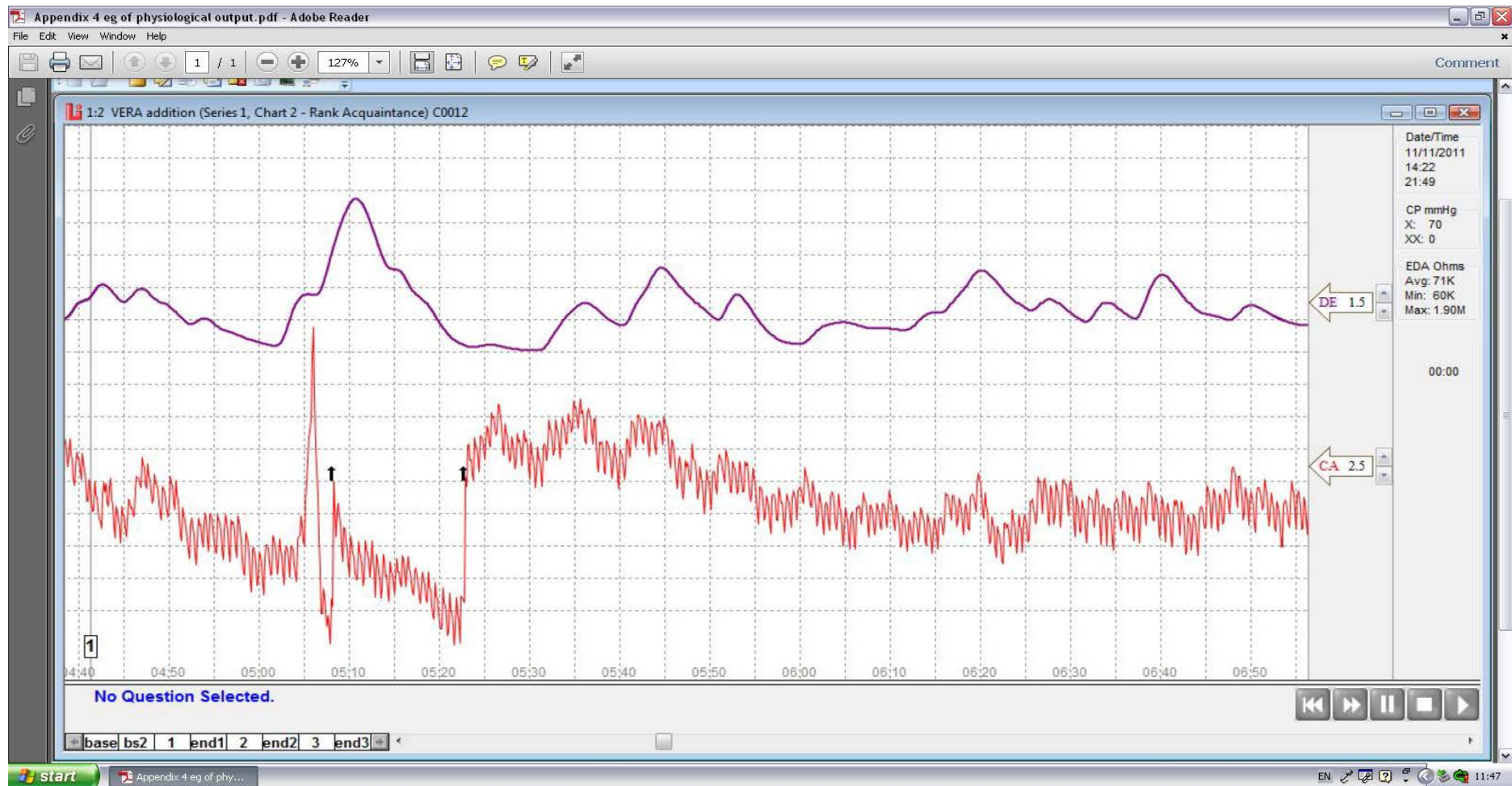
Continue

APPENDIX 5

When the arson happened, how did the person feel?		Not at all	A little	Somewhat	A lot	Very much so
1	How Worried did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	How Sad did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	How Upset inside did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	How Distressed did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	How Thrilled did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	How Sorry (regret that the incident happened) did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	How Disturbed did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	How Angry did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	How Interested did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	How Happy did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11	How Disgusted did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12	How Excited did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13	How Frightened did they feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Continue

APPENDIX 6



SERVICE EVALUATION PROJECT

Use of the Mental Health Clustering Tool in Croydon Perinatal Community Mental Health Team

Helen Lister

Institute of Psychiatry, King's College London

Supervised by: Dr Rachel Mycroft

CONTENTS

ABSTRACT	89
1. INTRODUCTION	90
1.1 Perinatal Mental Health	90
1.2 Croydon Perinatal Community Health Team (CMHT)	90
1.3 Kings Health Partners Clinical Academic Groups	91
1.4 Payment by Results	92
1.5 Mental Health Clustering Tool	93
1.6 Rationale	95
1.7 Aims	96
2. METHOD	97
2.1. Participants	97
2.2. Procedure	97
2.3. Approval of the Study	98
3. RESULTS	99
3.1. Adherence to MHCT	99
3.2. Severity of Client Problems	101
3.3. Care Clusters	102
3.4. Prophylactic Referrals & Referrals with Current Mental Health Problems	103
3.5. Adherence to Paired Outcome Measures (HoNOS)	104
3.6. Paired Outcome Measures: Change over Time	104
4. DISCUSSION	107
4.1. Adherence to MHCT	107
4.2. Severity of Client Problems	107
4.3. Care Clusters	108
4.4. Prophylactic Referrals & Referrals with Current Mental Health Problems	109
4.5. Adherence to Paired Outcome Measures (HoNOS)	110
4.6. Paired Outcome Measures: Change over Time	110
4.7. Limitations	111
5. CONCLUSION	113
6. REFERENCES	114
7. APPENDICES	117

LIST OF TABLES

Table 3.1. Demographic Information by Category	100
Table 3.2. Paired Outcome Measures (HoNOS) – Descriptive Statistics	105
Table 4.1. Care Clusters Identified	108

LIST OF FIGURES

Figure 3.1. MHCT: Mean Scores for Item Category (n = 55).....	101
Figure 3.2. Composition of Problems Identified for 'Item 8: Other'	102
Figure 3.3. Care Clusters Identified of 21 Possible Care Clusters.....	103
Figure 3.4. Care Clusters Identified Only.....	103
Figure 3.5. MHCT: Mean Scores for Item Category (n = 54).....	104
Figure 3.6. HoNOS Paired Measures.....	105

APPENDICES

Appendix 1. 21 Care Clusters.....	118
Appendix 2. Mental Health Clustering Tool (MHCT).....	119
Appendix 3. Cluster Decision Tree	126
Appendix 4. Graph: Distribution of Care Cluster Identified for Clients Per Caseholder.....	127

ABSTRACT

Mental health problems are relatively common during the perinatal period and have far-reaching consequences for women, their babies and their families. Investment in treatment during this period can prevent the development of long-term conditions in the mother whilst also 'inoculating' the infant against future mental health, cognitive and social difficulties. Croydon Perinatal Community Mental Health Team provides a specialist service with an emphasis on both short and long-term outcomes for mothers and their babies; also providing prophylactic services to women who are at high risk of relapse during the perinatal period. Lord Darzi (2008) stated that mental health currencies would be introduced in 2010/11 for the contracting and payment of mental health services and to facilitate benchmarking and comparison. To achieve this the Mental Health Clustering Tool (MHCT) was developed to allocate patients into Mental Health Care Clusters, from which specified care pathways will be provided. They will also inform the restructuration of SLaM into Mental Health Clinical Academic Groups (CAGs); new structures to bring clinical services and academic activities together within a series of single managerial units. Croydon Perinatal CMHT has been placed within the Psychological Medicine CAG (associated primarily with Care Clusters 7 and 14). The current study aimed to ascertain a baseline of the severity of the clients, and which care clusters they would be allocated to, on the caseload of Croydon Perinatal CMHT over a four week period, using the MHCT. Where paired HoNOS measures were identified these were analysed to identify change in client severity over time. Results showed that Croydon Perinatal CMHT are meeting SLaM targets for completion of MHCT and paired HoNOS measures. Analysis of measures showed no significant change in client severity over time. Care clusters identified ranged from Care Clusters 2-11 with no clients being allocated to Care Clusters 7 or 14. Limitations of the methodology of the current study are discussed and recommendations for the future are made.

INTRODUCTION

1.1. Perinatal Mental Health

Mental health problems are relatively common during the perinatal period. Severe disorders in particular can cause significant impact for women, their babies and their families during this time. It has been stated that women are at the highest risk of developing a serious mental illness in the first three months post-partum than at any other time in their lives (Austin 2004; 2010). Research by Kendal, Chalmers, and Platz (1987) provides evidence for this when stating that one to two women per thousand will develop an acute psychotic illness necessitating hospital admission in the early post-partum period. Community based research conducted by Thornicroft (1998) indicated that at least 2% of women who become pregnant and go on to have a child, have pre-existing, enduring and disabling mental health problems. Also, that they are at a substantially elevated risk of suffering a relapse or recurrence of a pre-existing disorder post-partum i.e. the rate of post partum relapse in women who suffer from bipolar affective disorder has been estimated at 33% to 50% (Marks, Wieck, Checkley, et al., 1992). This has far-reaching consequences including suicide, which is amongst the leading causes of maternal death (CEMACH, 2009), as well as adverse outcomes for the developing child i.e. post-partum mental illness is a significant factor in women who kill, injure or neglect their infants (Jayawant, Rawlinson, & Gibbon, 1998; Marks & Kumar, 1993) and is also associated with marital breakdown. Furthermore, antenatal and postnatal depression in mothers have been shown to have substantial and long-term deficits on the child's cognitive, emotional, behavioural and social development (Hay, Pawlby, Angold, Harold, & Sharp, 2003; Hay et al, 2001; Kim-Cohen et al., 2005).

The National Institute of Clinical Excellence (NICE) Guidelines for Antenatal and Postnatal Mental Health (2007) recommend rapid detection and treatment of mental health disorders in pregnant and postnatal women by specialist multidisciplinary perinatal mental health services. This is based on the notion that investment in treatment during the perinatal period can prevent the development of long-term conditions in the mother whilst also 'inoculating' the infant against future mental health and cognitive difficulties. This has been highlighted more recently in the Department of Health New Horizons document (2009) which states that mental health needs to be considered from birth onwards; from a life span perspective. Due to the negative effects of psychotropic medication on the foetus and breastfeeding infant, NICE recommends that the first line treatments during the perinatal period be psychological. Also, that a lower threshold be set for psychological intervention during this time and that mothers be seen quickly (normally within one month of initial assessment and no longer than 3 months afterwards).

1.2. Croydon Perinatal Community Mental Health Team (CMHT)

1.2.1. Purpose

To provide specialist assessment and treatment to women during their pregnancy, or up to 12 months postnatally, who are experiencing severe and enduring mental health problems and are

resident within the London Borough of Croydon. The London Borough of Croydon is densely populated, with a high birth rate and high levels of psychiatric morbidity. The annual birth rate for Croydon in 2009 was 5235 live births, the third highest of the Outer London boroughs. The fertility rate of 70 births per 1000 women aged 15-44 is above the national and the London average. (Office of National Statistics, 2010). Estimates of prevalence of mental health problems in the perinatal period predict that assuming 5235 live births in a twelve month period, 10% of these mothers (523) will experience postnatal depression, up to half of which (261) will be severe. The aim of the team is to promote both psychological well-being of mothers and positive relationships with their babies in order to safeguard the mental health of both the mother and her baby in both the short- and longer-term. The team also monitor and prophylactically treat women who are at high risk of relapse during the perinatal period.

Assessment and treatment are guided by the NICE Guidelines for Antenatal and Postnatal Mental Health and the SLaM Policy for the Care and Support of Pregnant Women with a Diagnosis of Severe Mental Illness. Treatment is adapted to take into account specific issues related to the puerperium, such as expectations and adjustment to the parental role and developing a positive relationship with the baby. The team also provide liaison with other services, both statutory and non-statutory, that work and provide services for women.

1.2.2. Team Composition

The team is comprised of multi-professional members of staff, all female and mostly part-time workers including: a team manager, a consultant psychiatrist, a junior doctor speciality trainee (GP), a senior clinical psychologist, three community psychiatric nurses (CPN's) and a team administrator.

1.2.3. Diagnosis of clients

An audit based on the team's caseload of 116 women on 23/04/2009 assessed the number of clients active to the team with specific diagnoses. Results showed that depression was the most common diagnosis of women on the team's caseload at this time, accounting for 62% (72 women).

1.3. King's Health Partners Clinical Academic Groups

King's Health Partners (KHP) is the collective name given to the Academic Health Sciences Centre (AHSC) established through the joining of Kings College London, Guy's and St Thomas' Hospital's, Kings College Hospital and South London and Maudsley NHS Foundation Trust (SLaM). The King's Health Partners mission is to integrate research, education, training and service delivery; breaking traditional organisational boundaries to allow professionals to work together across clinical services and academia to translate high quality research into practice more reliably, consistently and systematically. Clinical Academic Groups (CAGs) have been created in order to achieve this.

CAGs are new structures bringing clinical services and academic activities together within a series of single managerial units for the benefit of patient care. The aim of using CAGs are to provide more specialist services and more focused interventions, to improve consistency of care and greater clarification of what interventions are being provided and for whom within a particular disorder area. Key themes of CAGs include: care pathway development (including the introduction of mental health currencies in the form of payment by results and mental health clustering), promoting the integration of physical, psychological and social care, strengthening links between acute and mental health services, emphasising recovery and empowering teams to be innovative. To achieve this they are based on a multi-disciplinary approach to care and treatment.

At the time of writing, 21 CAGs have been identified, seven of which are mental health. The seven mental health CAGs are: Addictions; Behavioural and Developmental Psychiatry; Child and Adolescent Mental Health Services; Mental Health of Older Adults and Dementia; Mood, Anxiety and Personality Disorders; Psychological Medicine; and Psychosis. There is also the Clinical Neurosciences CAG that runs across KHP. SLAM aimed to implement the restructuration of its existing services into mental health CAGs by April 2010 to coincide with the introduction of mental health currencies.

1.4. Payment by Results

A commitment was made in Lord Darzi's *'High Quality Care for All'* (2008) that mental health 'currencies' would be available for use in 2010/11, where currencies are to be used for contracting and commissioning mental health services. The rationale for introducing currencies was to produce a consistent methodology for the contracting and payment for mental health services and to facilitate benchmarking and comparison; in short services will be paid according to who they treat and how successful their interventions are. To achieve this the Department of Health (DoH) introduced the concept of 'Payment by Results' (PbR) to mental health services (a strategy already successfully implemented into the acute healthcare sector).

PbR is the term given to the system for charging standard prices for standard units of activity within the NHS secondary care setting. As mentioned above payment is linked to activity and adjusted for case-mix. To implement currencies it has been necessary for services to allocate patients into distinct groupings or 'care clusters' based on the similar needs of patients, for whom specified care pathways will be provided. Within SLAM care clusters were established in conjunction with the restructuration of services into mental health CAGs. The activity delivered by services within these care pathways will be evaluated through the use of validated outcome measures.

An outcome measure provides a numerical estimate of the change that has taken place as a result of some sort of intervention or process. Clinically it is often used to mean "the positive changes, benefits, learning or other effects that result from the work that clinicians do" (Mental Health Outcomes Compendium, 2008). With regards to service development, outcome

measures should yield relevant information that is of value in making decisions about health care and help establish which interventions and services are desirable and cost effective. To enable patients to be allocated into 'care clusters' the Mental Health Clustering Tool (MHCT) was developed.

1.5. Mental Health Clustering Tool

The MHCT was developed by the Care Pathways and Packages Project to enable clinicians to allocate patients into 'care clusters' based on similar needs of patients. The MHCT incorporates the first 12 items from the Health of the Nations Outcome Scales (HoNOS) (Wing et al., 1999) as well as items from the Summary of Assessments of Risk and Need (SARN) (Self et al., 2008), in order to provide all the information necessary to allocate individuals to clusters.

In total the MHCT is an 18 item tool, essentially the HoNOS with the addition of 6 further scales measuring beliefs in non-psychotic patients, risk, vulnerability and engagement, see Appendix 1.

- **PART 1** contains 13 scales to record problems experienced during the 2 week period prior to assessment (the current rating period)
- **PART 2** contains another five scales (A-E) that consider problems from a 'historical' perspective. These will be problems that occur in episodic or unpredictable ways. While they may not have been experienced by the individual during the 2 week current rating period, clinical judgement would suggest there is still a cause for concern that cannot be disregarded (i.e. no evidence to suggest that the person has changed since the last occurrence either as a result of time, therapy, medication or environment etc.) In these circumstances any event that remains relevant to the cluster allocation (and hence the interventions offered) should be included

All scales are rated as for the original HoNOS. Patients are allocated into care clusters dependent on the scores that are evidenced.

Based on the information gathered during the routine screening / assessment process, the clinician scores the service users' needs using the MHCT. The clinician then decides if the origin of the needs is primarily non-psychotic, psychotic or organic, and then narrows the list down to the possible Mental Health Care Clusters by identifying the most accurate sub-category. The clinician finally uses the MHCT and their clinical judgement (and / or consults an online clustering algorithm) to decide which care cluster is the most appropriate for the patient. At the time of writing an algorithm was being created to be used within SLAM which will electronically calculate the most appropriate care cluster for the client. Once completed it will be used to aid the clinician in their decision. However, if the clinician disagrees with the projected care cluster he/she can override this manually, stating their reasons for doing so.

SLaM recommends that the MHCT be completed at key time points in the patients journey including: referral to the team or service; planned reviews (Care Programme Approach (CPA)/6 month); crisis (or any changes in planned care/need) and discharge or transfer from a team or service i.e. at any point where a significant change in need occurs. The minimum requirements for HoNOS however are admission, discharge and 6 monthly review points. Only the first 12 original HoNOS items of the MHCT are used to measure health outcome and these items can also be used at any other time clinicians deem appropriate to monitor the impact of specific clinical interventions. An HoNOS score can therefore be extracted from the MHCT using the total of the first 12 items.

1.5.1. HoNOS

HoNOS is an integral part of the Mental Health Minimum Data Set, (it has been part of the MHMDS within SLaM since 2005). It is an internationally recognised outcome measure developed by the Royal College of Psychiatrists Research Unit (RCPRU), to be completed routinely by clinicians and recorded as part of the MHMDS. It measures the broad range of physical, psychological and social problems associated with severe mental illness. It contains 12 scales / items. The 12 scale items provide a summary of the key problems which affect people with severe mental illness. The key problems can be grouped into 4 categories: Behavioural Problems (consisting of items 1-3; Impairment (items 4 and 5); Problem Symptoms (items 6-8) and Social Problems (including items 9-12). Each item is rated using a 5 point scale from 0-4, where 0 is “no problem” and 4 is “severe / very severe problem”, with overall scores ranging from 0–48, where 0 is “no distress” and 48 is “highest distress”. A score of 1 on each item is classified as being of sub-clinical severity. It is not an interview but a clinical assessment based on information collected during a defined period of time. Wing et al. (1998) have provided evidence to show that it is reliable, clinically useful, acceptable, sensitive to change and useful for administration and planning in both in-patient and community settings. This has been supported in numerous studies to date; Orrell, Yard, Handysides, and Schapira (1999); Amin, Singh, Croudace, et al. (1999); Slade, Thornicroft, Beck, Bindman, and Wright (2000); Rees, Richard, and Sharp (2004) and Parabiaghi, Barbato, D’avanzo, Erlicher, and Lora (2005).

A single HoNOS provides a profile of “present state” and a measurement of severity. Two or more HoNOS records completed during an episode of care / treatment provide an estimate of health outcome associated with clinical activity and the effectiveness of clinical services. Outcome is measured by comparing individual items scores, subscale scores or total scores at different time periods. Serial ratings allow the measurement of health status and problem trends and fluctuations over time. It is for these reasons that it has been adopted for use with major CPA reviews (Audin, Margison, Mellor Clark, & Barkham, 2001) and indeed as part of the MHMDS by commissioners of NHS Trusts, including SLaM. At the time of data collection, SLaM currently employed a target of 60% HoNOS paired measures for clients on clinician caseloads that are eligible. ‘Paired measures’ defines two HoNOS completed for one patient over a period of time. The period of time can be defined as assessment and again at discharge,

or clients who have an initial HoNOS at assessment and are active to services for 6 months or longer and therefore require HoNOS to be repeated on a 6 monthly basis. Also at assessment and when another measure has been conducted by the clinician wanting to monitor the effect of a specific intervention. Currently SLAM has a target for completion of paired HoNOS measures of 60%.

1.5.2. Care Clusters

Currently SLAM have identified 21 Mental Health Care Clusters within the mental health CAGs that together form the 'currencies' for contracting their mental health services, see Appendix 2. (This may increase once the clusters have been evaluated; based on statistical analysis and feedback from their clinical use).

Croydon Perinatal CMHT has been placed within the Psychological Medicine CAG, see Appendix 3. It has been predicted that patients needs within the Psychological Medicine CAG will be best served within Care Cluster 7: Enduring Non-Psychotic Disorders (High Disability) and Care Cluster 14: Psychotic Crisis.

1.6. Rationale

The requirement to allocate all service users seen during the financial year 2010-2011 into mental health care clusters has been incorporated into the Commissioning for Quality and Innovation (CQUIN) indicators for London Mental Health Trusts. The CQUIN framework apportions an element of the provider's income based on quality and innovation targets. Where SLAM can meet specified targets for clustering their patients they can receive extra funding for their services in 2011, including Croydon Perinatal CMHT. Consequently SLAM has set completion targets for clinicians to allocate their patients to clusters; such that upon implementation of the MHCT all new clients accepted by services are to be allocated to care clusters by September 2010. This is in addition to achieving existing targets as mentioned earlier, as part of the MHMDS, for clinicians to complete paired outcome measures (HoNOS) - set at 60%. Data from the completion of these outcome measures can then be used to ascertain the severity of the clients on the caseload of Croydon Perinatal CMHT, among other services. Also, if there are sufficient paired measures then this data can be used to assess statistically reliable and clinically significant change over time of the clients on the caseload.

Outcome data from the MHCT can be used to ascertain a baseline of the care clusters that Croydon Perinatal CMHT practitioners have allocated their current caseload to. This would provide some preliminary information as to the composition of the team's current caseload, which will provide the team with information on how to develop their service, based on the range of client needs identified for those on the current caseload.

Women with a history of bipolar disorder, postpartum psychosis or other psychosis are at high risk of relapse during the perinatal period and are therefore seen by the Perinatal Team for monitoring, preventative work and treatment as required. This lower threshold for intervention

for disorders occurring during the perinatal period is in accordance with NICE Guidelines for Antenatal and Postnatal Mental Health (2007). However as these women are currently well at referral this may affect the severity of the client caseload mix and indeed the composition of care clusters of clients on the team caseload.

1.7. Aims

- 1) To assess adherence to completion of the MHCT by caseholders of Croydon Perinatal CMHT in line with SLaM targets
- 2) To assess the severity of clients on the caseload of Croydon Perinatal CMHT using the HoNOS scores / HoNOS 12 items score of MHCT
- 3) To identify a baseline of which care clusters the clients on the caseload of Croydon Perinatal CMHT have been allocated to
- 4) To compare the severity of prophylactically referred clients with those referred with current mental health problems
- 5) To assess adherence to completion of paired outcome measures (HoNOS), set at 60% of the caseload
- 6) If there are sufficient HoNOS paired measures, to assess for statistically reliable and clinically significant change for the clients on the caseload of Croydon Perinatal CMHT

2. METHOD

2.1. Participants

The participants consisted of all 89 patients active to the caseload of Croydon Perinatal CMHT as of 17th May 2010.

2.2. Procedure

At the time of the project, the MHCT had only just been introduced within SLaM therefore all caseholders of the Croydon Perinatal CMHT attended a SLaM training day entitled “The Mental Health Clustering Tool”. This was facilitated by Kevin Smith (Trust Clinical Outcomes Team). The training day gave an introduction to the tool and also included practical examples of how to complete the tool through the use of group discussions. Once staff had attended this training day they were deemed competent to use the MHCT.

During a nominated period of four weeks (17th May to 13th June 2010) caseholders of Croydon Perinatal CMHT completed as many MHCTs as they could for the clients that they saw within this period, they entered these scores into individual client files on SLaM's electronic patient journey system (ePJS) then allocated individuals to care clusters based on this information. During the planning stage of the research the MHCT was still under development. However, once data collection was underway SLaM had implemented the tool. This also included successfully updating ePJS to include the measure electronically, meaning that the caseholders were able to enter their scores directly on to ePJS. At this time a clustering algorithm, designed to electronically allocate clients to care clusters was still not active therefore caseholders manually allocated their clients to care clusters. A total caseload list, as determined on 17th May 2010, was provided to each caseholder from the researcher in order to assist with this. The total team caseload as determined on 17th May 2010 was 89. Each caseholder was then requested to transfer this data over to a Microsoft Excel spreadsheet as provided by the researcher. This was to allow the researcher to keep track of the progress of the individual caseholders to complete the MHCT for their clients as well as to allow the researcher to provide weekly prompts of which clients the caseholders had yet to complete the tool for. Additional data was also extracted from ePJS by the Trust Outcomes Team. This data was used by the researcher to provide detailed demographic data of the sample as well as to compare with the data that had been manually entered by the caseholders.

Prophylactic referrals were distinguished from referrals for clients who were currently experiencing mental health problems by accessing diagnostic codes attributed to each client when the referral was accepted by the team, from a Microsoft Excel spreadsheet held by the team administrator. Prophylactic referrals were coded as Z71.1 and all other codes were combined to form the group of ‘referrals currently experiencing mental health problems’.

2.3. Approval of the Study

The study was registered with the Clinical Governance & Audit Department of South London and Maudsley NHS Foundation Trust.

3. RESULTS

3.1. Adherence to MHCT

During the nominated four week period 58 of 89 (65%) of the clients on the caseload identified on 17th May 2010 were seen by the caseholders of Croydon Perinatal CMHT. Of these, 55 of 58 seen (95%), were allocated to care clusters.

3.1.1. Demographics

All clients were female and residents of the London Borough of Croydon. The mean age of the 55 clients who were seen and allocated to care clusters was 31 years with the range from 19 years to 48 years. The majority of the 55 clients seen and allocated to care clusters cited their marital status as single or married / civil partnership, accounting for 48% and 45% respectively. The majority of clients 33 (60%) stated their ethnicity as British.

Table 3.1. shows a more specific breakdown of demographic information by category.

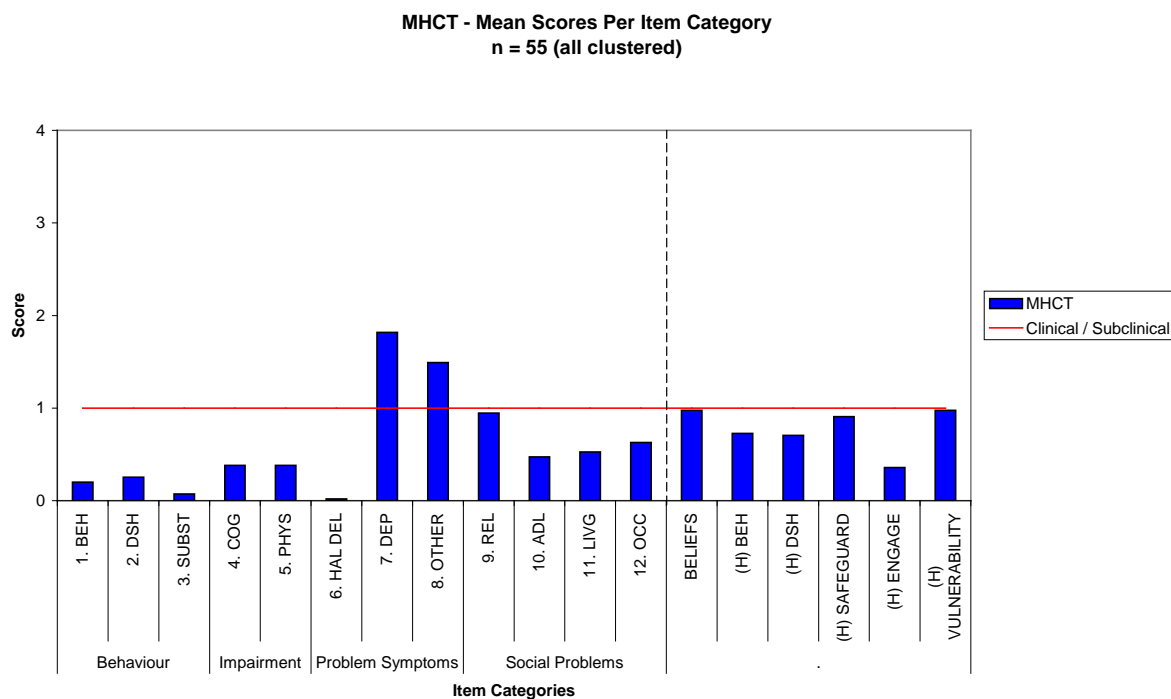
Table 3.1. Demographic Information by Category

Category		Number of Clients	Percentage
Age Group:	17-25	10	18
	26-34	27	49
	35-44	17	31
	45-54	1	2
Marital Status:	Single	26	48
	Married / Civil Partnership	25	45
	Separated	3	5
	Co-habiting	1	2
Ethnicity:	British A	33	60
	Caribbean (M)	3	5
	Ugandan (PQ)	3	5
	Any Other Group (SE)	2	4
	Other White Unspecified (C3)	2	4
	Black British (PD)	1	2
	British Asian (LH)	1	2
	English (CA)	1	2
	Indian/British Indian (H)	1	2
	Iraqi (SJ)	1	2
	Irish (B)	1	2
	Middle Eastern (SF)	1	2
	Nigerian (PC)	1	2
	Other African (N)	1	2
	Other Asian Unspecified (LK)	1	2
	Pakistani/British Pakistani (J)	1	2
	Scottish (CB)	1	2

3.2. Severity of Client Problems

The distribution of client problems across the item categories of the MHCT and the mean severity of client problems are shown in Figure 3.1.

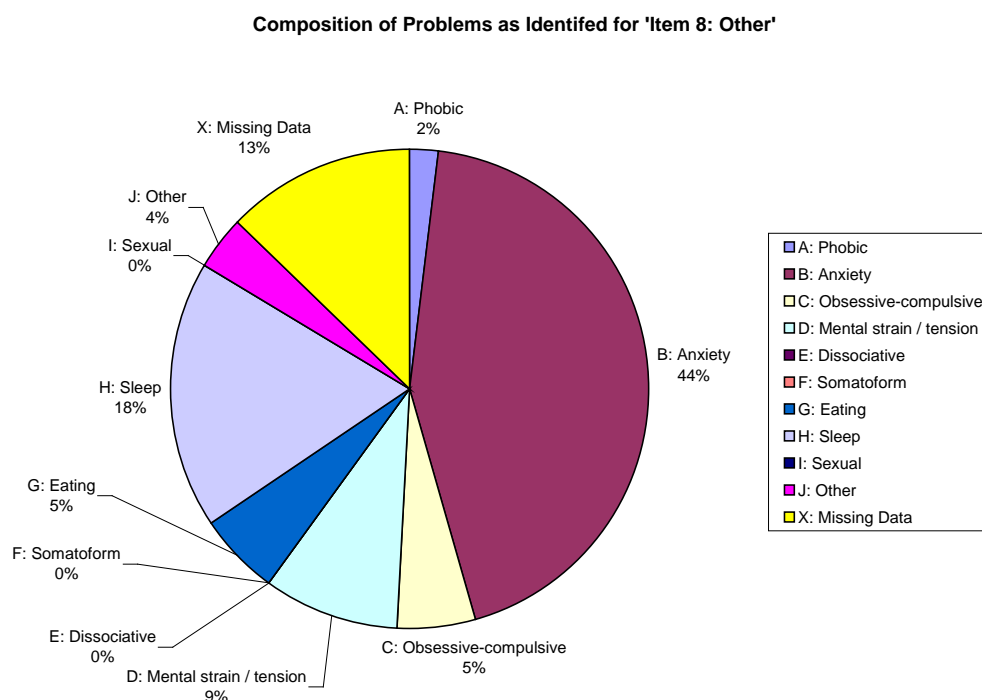
Figure 3.1.



As Figure 3.1. shows 'Item 7: Depression' and 'Item 8: Other' were the most severe problems identified for the clients of Croydon Perinatal CMHT during the period of data collection. Both items are above the clinical cut-off of 1 with mean scores of 1.82 and 1.49 respectively. 'Item 9: Relationships' was identified as having the third highest mean score at 0.95. However, this item remains within the subclinical level as it is less than the clinical cut-off of 1.

Figure 3.2. shows the composition of problems as identified by Item: 8 'Other'.

Figure 3.2.



The most commonly occurring problem experienced by clients, during the data collection period, within the category of 'Other' was Anxiety; accounting for 51%. This was followed by problems with Sleep (18%) and Mental State / Tension (9%). Missing data from this category accounted for 13%.

3.3. Care Clusters

Overall, clients were allocated to care clusters ranging from care cluster 2- 11, the majority of which being allocated to care clusters 3 and 4.

Figures 3.3. and 3.4. show the composition of care clusters belonging to the 58 clients who were seen and allocated to care clusters during the data collection period. Figure 3.3. shows this in relation to all 21 possible care clusters. Figure 3.4. shows only the care clusters that were identified, where purple represents non-psychotic care clusters; blue represents psychotic care clusters and yellow represents substance misuse / clients not allocated to a care cluster.

Figure 3.3.

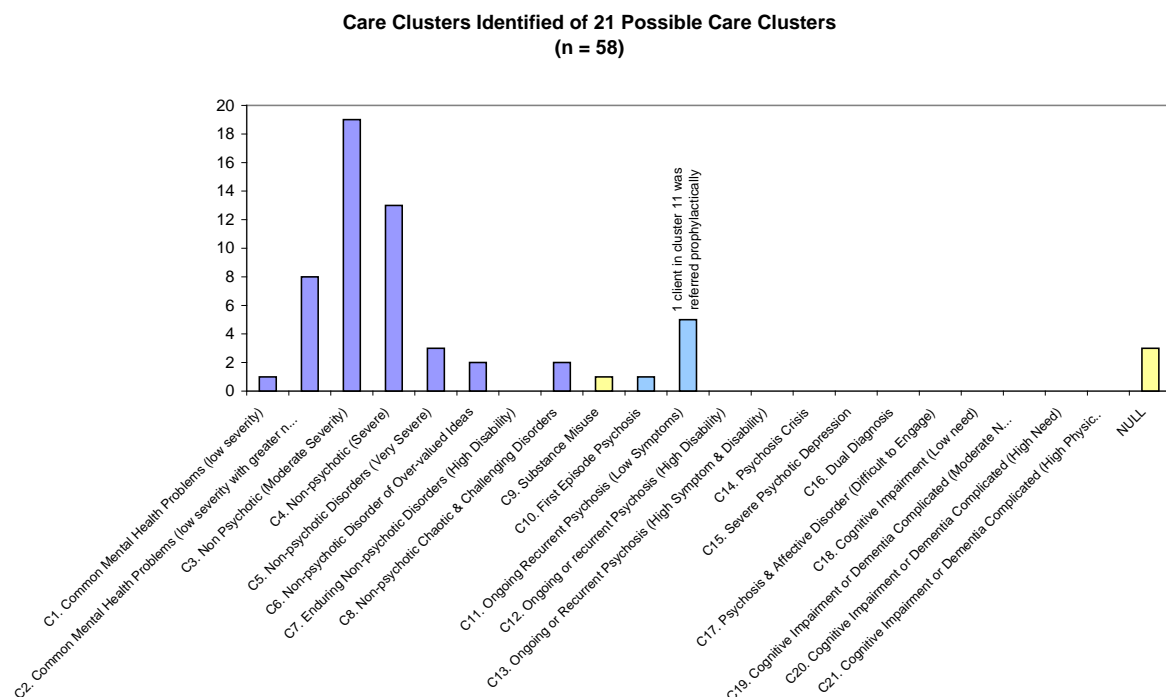
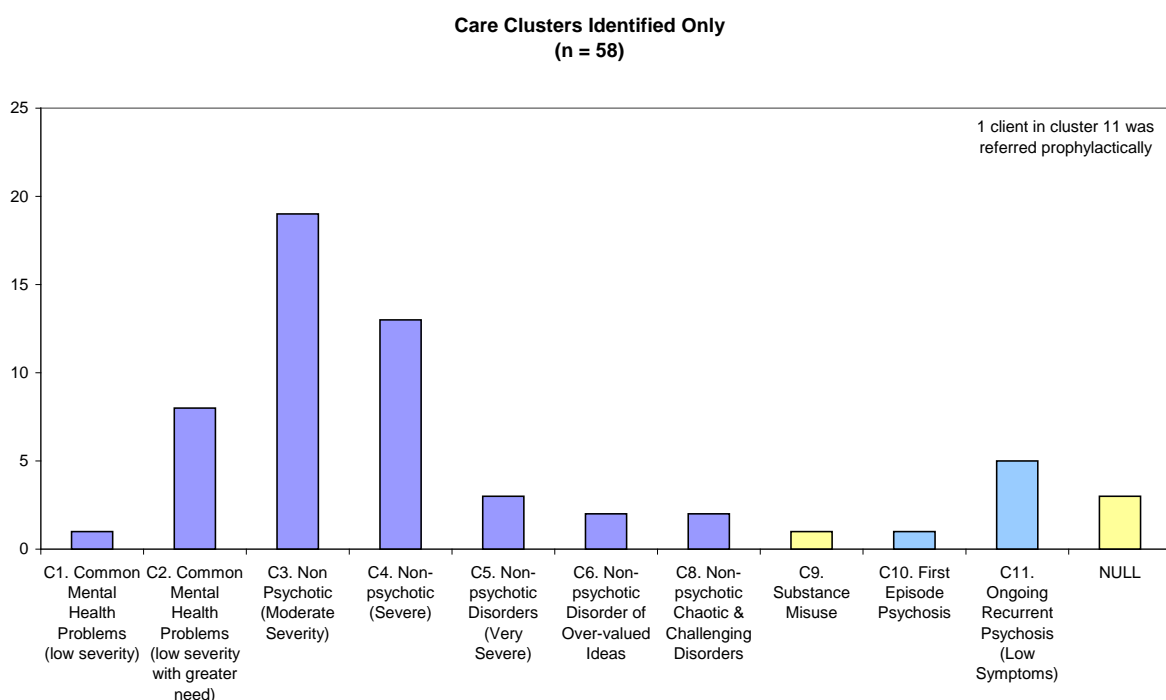


Figure 3.4.



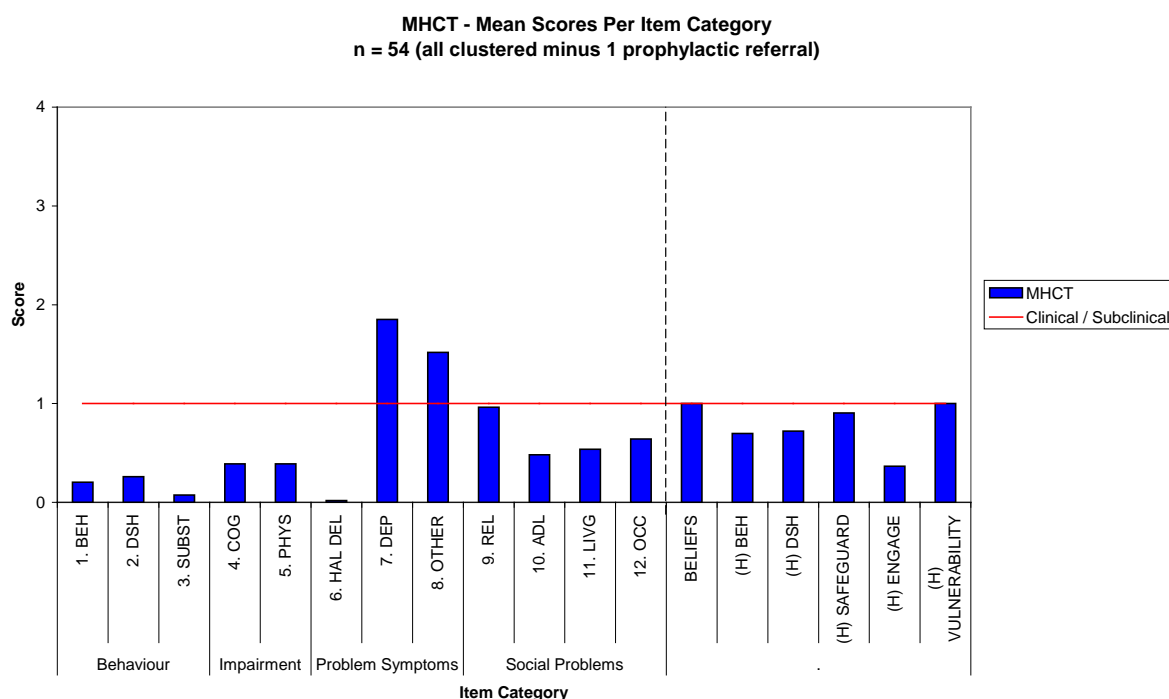
3.4. Prophylactic Referrals and Referrals with Current Mental Health Problems

Overall of the 89 clients identified on the caseload on 17th May 2010, only two were identified as being prophylactic referrals. Of the 55 clients seen and allocated to care clusters, only one was identified to be a prophylactic referral therefore due to the small sample size, resulting in the

assumption of homogeneity of variance being unmet, any further inferential analysis between this population and those referred with current mental health problems was not conducted.

Figure 3.5. shows the distribution of client problems across the item categories of the MHCT and the mean severity of client problems minus the prophylactic referral.

Figure 3.5.



The care cluster that this client was allocated to was 'Cluster 11: Ongoing Recurrent Psychosis (Low Symptoms)', as shown in Figures 3.3. and 3.4.

During the data collection period this client was identified as having paired outcome measures, recording a total HoNOS score of 0 for both first and most recent HoNOS measures.

3.5. Adherence to Paired Outcome Measures (HoNOS)

Of the 55 clients who were seen and allocated to care clusters, 44 (80%) had paired HoNOS measures (including original HoNOS scores and/ or HoNOS scores derived from the first 12 items of the MHCT). This gives an adherence rate of 80% paired HoNOS measures.

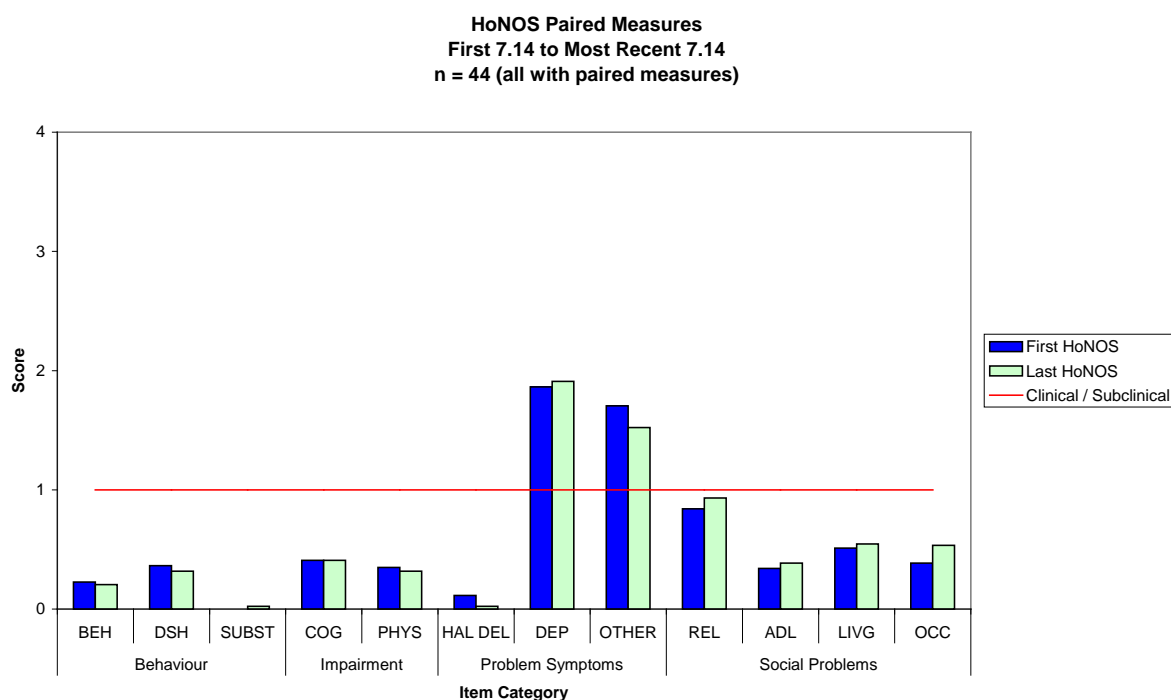
3.6. Paired Outcome Measures: Change over Time

3.6.1. Descriptive Statistics

Preliminary descriptive statistics shown in Table 3.2. suggest that there was no change in mean severity over time, represented graphically in Figure 3.6.

Table 3.2. Paired Outcome Measures (HoNOS) – Descriptive Statistics

	First HoNOS Score	Most Recent HoNOS Score
Average	7.31	7.31
Standard Deviation	4.95	5.51
Range	0 – 24	0 – 23

Figure 3.6.

As figure 3.6. shows there appears to be a rise in the mean score for 'Item 7: Depression'; rising from 1.86 to 1.91. The mean score for 'Item 8: Other' has reduced from 1.70 to 1.52 although remains above the clinical cut-off of 1. 'Item 9: Relationships'; 'Item 10: Activities of Daily Living', 'Item 11: Living Conditions' and 'Item 12: Occupation and Activities' have all risen also. There were reductions in mean severity scores of 'Item 1: Behaviour', 'Item 2: Non-accidental Self Injury', 'Item 5: Physical illness or Disability Problems' and 'Item 6: Problems with hallucinations and Delusions'. There was no change with 'Item 4: Cognitive Problems'.

Despite preliminary descriptive results suggesting that there was no change in mean severity scores over time further inferential analysis was conducted to confirm this. The data were analysed using SPSS version 15.0.

3.6.2. Statistically Significant Change

A repeated measures t-test was employed to examine if there was any statistically significant difference in mean severity scores over time, thus to determine if there was any change in the severity of clients problems over time (of those clients who were seen and allocated to care clusters and who had paired HoNOS measures; the one prophylactic client was removed from

this analysis therefore $n = 43$). Results of the t-test showed that there was no statistically significant difference over time;

$$t(42) = 0.006; p > 0.05 \text{ (not sig.)}$$

3.6.3. Reliable and Clinically Significant Change

Parabiaghi, Barbato, D'avanzo, Erlicher, and Lora (2005) presented a methodological framework for calculating reliable and clinically significant change using HoNOS, based on the Jacobson and Truax (1991) model. However, as the current data showed no statistically significant change in severity over time, further calculations based on this model were not undertaken.

4. DISCUSSION

The results presented above are discussed in reference to the six original aims of this project. Limitations of the study and implications for future practice are discussed before conclusions are drawn.

4.1. Adherence to MHCT

During the specified four week data collection period set by the researcher, the caseholders of Croydon Perinatal CMHT achieved an adherence rate of 95% for completion of the MHCT for the clients on the team's caseload. They have demonstrated that they are adhering and working towards meeting the objective set out by SLAM in relation to the MHCT, which states that all new referrals to the team are to have the MHCT completed by September 2010. They have achieved this despite also working in retrospect (completing the tool for current patients on the caseload identified during the data collection period) as well as any new clients seen during this time, for the purposes of the project. This is an excellent achievement for the service and highlights the great dedication of the team to implement the measure within the service, both to allow for a detailed picture of the composition of the perinatal population within SLAM as well as to benefit service delivery.

4.2. Severity of Client Problems

As the results of the project have shown the main problems identified of the population treated in the perinatal period by Croydon Perinatal CMHT are 'Depression' and 'Other'; items 7 and 8 of the HoNOS / first 12 items of the MHCT, respectively. The mean severity of these problems for the population both fell above the clinical cut-off of 1, at 1.82 and 1.49, above that of the mean severity of all other items measured by the HoNOS, which were within sub-clinical levels. An HoNOS score of 1 for a client is described as a 'minor problem' and 2 as a 'mild problem' therefore it can be concluded from these results that the overall severity of the problems identified for this population is relatively low. This suggests that the team are adhering to the NICE guidelines for Antenatal and Postnatal Mental Health (2007) by setting a lower threshold for the perinatal population to be accepted for intervention.

The finding that depression is the main problem of the caseload studied in the current project is consistent with those of a previous audit of the team's caseload in 2009, which found the most common diagnosis of women at that time was also depression; a diagnosis for which NICE recommends psychological therapy. 'Item 8: Other' when analysed into its component parts, reveals anxiety to be the most commonly occurring problem within this category, accounting for 44% of the variance. Again, a problem which is often found co-morbid with depression (point-prevalence figures from the National Psychiatric Morbidity Study – UK show mixed anxiety and depression to account for 7.7% of the population) and for which NICE recommends psychological therapy.

The items of the HoNOS (first 12 items of the MHCT) can be further grouped into four categories: Behavioural Problems; Impairment; Problem Symptoms and Social Problems. When applying these categories to the results of the current study it is evident that the clients of Croydon Perinatal CMHT have the most severe difficulties within the Problem Symptoms and Social Symptoms categories than the Behaviour and Impairment categories.

It is within the Problem Symptoms category where clients are within the clinical range for depression and anxiety, as would be expected for clients to be accepted by the service for treatment. However, looking closely within the Social Symptoms category provides evidence that although these component items ('Relationships', 'Occupation and Activities', 'Living Conditions' and 'Activities of Daily Living') are not causing difficulties within clinical levels, with mean severity scores of 0.95, 0.63, 0.53 and 0.47 respectively, these 'social problems' may also be contributory factors to the main problems identified of depression and anxiety symptoms, thereby increasing the severity of the latter symptoms to within clinical levels. In short, this suggests that although the Social Symptoms for the perinatal population are below the clinical cut-off they may be additional factors exacerbating the severity of the Problems Symptoms category. This finding provides support for the treatment aims of the team in that not only is treatment adapted to take into account specific issues related to the puerperium, but the team also provide liaison with other services, both statutory and non-statutory that provide services for women for example, Relate and Housing Associations, among others.

Qualitative information gathered from the team following discussion of these findings provides evidence which suggests that this is an accurate reflection of their clients' problems. Caseholders stated that a lot of their clients are struggling with maintaining relationships and attaining appropriate housing / living conditions for their family. For example, some clients may be looking for larger housing to home their expanding family, some clients may be experiencing problems with regards to maternity leave i.e. when to return back to work, or indeed the difficult emotions that are evoked when a mother either makes the transition into stay-at-home mum or the transition of returning to employment after maternity leave, and the associated difficulties that childcare arrangements can generate etc. Furthermore, Croydon is known as an over-populated and somewhat deprived borough with a high unemployment rate and therefore problems with adequate housing and employment are high for a number of Croydon residents, which may be exacerbated by pregnancy.

4.3. Care Clusters

The main care clusters identified from the population are included in Table 4.1.

Table 4.1. Care Clusters Identified

Cluster Number	Cluster Name	Number of Clients in Cluster
Cluster 3:	Non Psychotic (Moderate Severity)	19
Cluster 4:	Non Psychotic (Severe)	13
Cluster 2:	Common Mental Health Problems (low severity with greater need)	8

As shown by the Care Clusters Decision Tree (Appendix 3) the three main care clusters identified within the current study would be placed within the Mood, Anxiety and Personality Disorder CAG and not within the Psychological Medicine CAG, which has been highlighted to serve the population of Croydon Perinatal CMHT. Furthermore as shown in Figure 3.4., no clients from the perinatal CMHT who were seen and allocated to care clusters during the data collection period were allocated to Cluster 7: Enduring Non-Psychotic Disorders (High Disability) or Cluster 14: Psychotic Crisis; the clusters that make up the Psychological Medicine CAG and which constitute the allocation of Croydon Perinatal CMHT within this CAG. This finding, in addition to the relatively low severity of the problems associated with the population treated in the perinatal period, suggests that the perinatal population may differ from other populations and therefore relate to care clusters different than those seen by other teams within the Psychological Medicine CAG. This suggests that this population may require special consideration i.e. setting lower thresholds for intervention and evaluation of treatment outcomes. It is suggested that the care clusters primarily identified within the Psychological Medicine CAG i.e. care clusters 7 and 14, may need refining.

4.4. Prophylactic Referrals and Referrals with Current Mental Health Problems

From the 89 clients identified as comprising the team caseload on 17 May 2010 only 2 (2%) of these were classified as prophylactic referrals. Of the 55 clients who were seen by the clinicians during the data collection period only 1 (2%) of these clients included a prophylactic referral, therefore of the identified caseload only one prophylactic referral was seen during the data collection period. As a result of there being only one prophylactic referral, thus there being only one client in this group, no inferential analysis was undertaken as the assumptions for this to be conducted were not met. One possible explanation to account for the very small population in this arm of the study may have been that although the team do see prophylactic clients this is not as regular as clinicians may perceive, possibly as a consequence of the relatively low mean severity of the problems of the population on the teams caseload. Alternatively, this study only looked at the teams caseload over a short period of time (four of fifty-two weeks), which may not have allowed an accurate representation of the number of prophylactic referrals to the service overall to be attained, for example it is possible that the service may receive more prophylactic referrals at different times of the year. Additionally, by definition prophylactic referrals are deemed to be currently well therefore as a consequence of the limited time period of the project clinicians may have delayed assessing these clients compared to clients referred with current mental health problems. In summary if the project was to be conducted again a longer period of time, or a different time of year may be chosen for data collection. Finally, for the purposes of the project prophylactic referrals were identified as those who did not otherwise have a formal current mental health diagnosis at the time of the referral being accepted by the team, which may have led to an under-estimation of this population. To address this issue, in consideration of future audits in this area, it may be useful to implement more sophisticated ways of identifying prophylactic referrals i.e. looking more specifically at ICD-10 codes, for example codes for history of bipolar disorder, currently in remission.

4.5. Adherence to Paired Outcome Measures (HoNOS)

During the data collection period the team achieved an adherence rate of 80% for paired outcome measures, which exceeds the 60% target stipulated by SLAM. Again, this is an excellent achievement by the team and suggests that prior to the project the team were achieving targets with relation to adherence to routine outcome measure completion.

4.6. Paired Outcome Measures: Change over Time

Results of the project, comparing the first and most recent HoNOS scores (first 12 items of the MHCT) to assess for change in severity over time found no statistically significant change in mean severity of client problems over time. As a result, in addition to the relatively small sample size of the study, no tests for reliable clinical change were conducted. This finding was somewhat unexpected following a review of the literature which highlighted numerous published studies where statistically reliable and reliable clinical change have been demonstrated using paired HoNOS measures with a clinical population e.g. Parabiaghi et al. (2005) and Jacobson and Truax (1991). There are two possible explanations to account for this finding.

Firstly, it is proposed that this may have occurred as a result of the variety of time periods between the first and the most recent HoNOS of the current project. In published studies time periods for paired HoNOS measures have typically been taken at assessment then at intervals no shorter than six months, or else at discharge. This is in contrast to the current project where paired measures were taken to be the first and most recent HoNOS for each client; where the time between measures was not controlled as a variable, due to the limited time span of the study. Consequently, for some clients the time period between their first and most recent HoNOS may have been considerably less than six months, in some cases this may have been as short as one week, for example, where the clinician had recently completed an HoNOS for a client but then also completed the MHCT (when this was introduced to enable the client to be allocated to a care cluster) the clinician may have simply duplicated the scores from the HoNOS to the MHCT. For the current project this would give the client paired HoNOS measures; the original HoNOS being considered the first HoNOS and the MHCT (first 12 items) being considered the most recent HoNOS. In addition, the vast majority of clients would not have completed their treatment at the time of the most recent HoNOS and therefore would be unlikely to demonstrate as much change on this measure. As such, comparisons between the findings of the current study and published studies for reliable and clinical change are not possible. It is hypothesised that had the current project used HoNOS at discharge instead of most recent HoNOS, the results may have differed.

Secondly, it is possible that some clients may have had up to six months or more between their scores with no change in the severity of their total HoNOS scores over the two time periods. When reviewing literature on the perinatal population this may not be wholly unexpected for example, it has been highlighted that mental health problems typically worsen during the perinatal period (NICE, 2007) therefore keeping women stable over this time may demonstrate

an effective service. Applying this hypothesis to the paired HoNOS data collated from the current project suggests that Croydon Perinatal CMHT are providing an effective service to the perinatal population; they are accepting referrals at a lower threshold of need for mental health problems (identified through first HoNOS) and through effective intervention over the perinatal period where mental health typically worsens, are maintaining the level of severity of their clients problems (most recent HoNOS).

However, these explanations are merely speculative therefore if the study was to be completed again, more stringent criteria would be employed of the difference in time periods spanning the paired HoNOS measures i.e. only including data where a six month interval between repeat HoNOS measures are identified, or else using the clients first HoNOS and discharge HoNOS instead of most recent HoNOS, to allow for comparisons to be made.

4.7. Limitations

The current project has a number of limitations, the main being the limited time span in which it was conducted. A fixed four week period was applied as this was felt to be the optimum time to request that the team participate in the study, in addition to their routine clinical duties. However, in doing so it was not possible to audit the entire caseload of the team. Consequently, it is possible that the sample obtained is not completely representative of the perinatal population, including the prophylactic population as discussed previously. Additionally, as mentioned, the limited data collection period was a factor in the decision not to control for time between paired HoNOS measures which, as discussed, had serious implications for establishing significant and reliable change in severity of client problems over time. Had the duration of the project been increased this would have allowed for a larger sample size thus making any effects found to be determined with certainty, which is not possible with the present data. Taken as a whole this has the effect of limiting the generalisability of the findings to the wider perinatal population.

Another limitation is that the MHCT had only just been introduced within SLAM at the time of the project. This resulted in a number of implications which may have affected the findings of the project for example, clinician confidence with the tool, as well as the tool being fully accessible electronically. To explain, although the team had recently completed training in the MHCT prior to the project beginning their confidence in using this relatively new outcome measure may not have been as high as if this was an established routine clinical outcome measure. Any uncertainty in completing the measure may have been exacerbated by clinicians having to manually allocate their clients to care clusters due to the unavailability of the electronic clustering algorithm. Combined, this may have affected the distribution of the care clusters to which clients were allocated, and may account for the 5% of cases that were not clustered.

Another consequence of clinicians having to manually allocate clients to care clusters, which may account for the range of care clusters identified, may be attributed to the individual variation of clinicians when manually clustering. The composition of care clusters per clinician's

caseload was identified (despite not being an original objective of the study) to highlight this, see Appendix 4. For example, it is interesting that one clinician clustered all clients into only one cluster compared to the rest of the team who had identified a range of clusters for their individual caseloads. It would be interesting to evaluate this further; comparing clinician care cluster allocation to those identified using the electronic algorithm. It must be noted however that there is also individual variation of clients and no inter-rater reliability of allocation to care clusters was conducted, which is a further limitation within the project and would be included if this was to be repeated again. In the future however, the clinician may be better guided as to which cluster to allocate clients to from the electronic clustering algorithm which may help to standardise the range of care clusters.

A similar limitation to the recent introduction of the MHCT within SLaM at the time of the project is the recent establishment of CAGs and care clusters within KHP. One of the aims of the current project was to establish a baseline of care clusters that clients treated by Croydon Perinatal CMHT were allocated to. This was achieved, however, care clusters 7 and 14 were not identified for this population as predicted by SLaM for the Psychological Medicine CAG which, although interesting, nevertheless raises questions regarding the reliability of care cluster allocation and CAGs. In conclusion to this limitation it is suggested that the findings of the current project be used as pilot data within SLaM to aid evaluation of the effectiveness of the allocation of the existing care clusters to their respective CAGs.

5. CONCLUSION

This project has found that over a four week period, Croydon Perinatal CMHT were meeting SLaM targets for completion of the newly implemented MHCT. The overall severity of the population treated during this period was low with only two of the twelve items of the HoNOS, 'Depression' and 'Other', identified as falling within the clinical range. A baseline of the care clusters that clients were allocated to was identified ranging from care clusters 2 – 11, the majority of which being allocated to care clusters 3 and 4. No clients were allocated to care clusters 7 and 14 of the Psychological Medicine CAG as expected. Only one of two prophylactic referrals identified was seen during the data collection period therefore it was not possible to conduct a comparison between the severity of this population and those on the caseload identified as experiencing current mental health problems. SLaM targets for completion of paired outcome measures (HoNOS) was achieved and exceeded by the team during the data collection period at 80%, analysis of which found no significant change in severity of problems over time. Limitations include, among others, the relatively short time span of the project which restricted the possible sample size thus limiting the representativeness of the population. In addition, the time from first to most recent HoNOS was not recorded or controlled for. Nevertheless this project could act as a guide to a larger audit of the service. Overall, it is concluded that the problems associated with the population treated in the perinatal period require special consideration, as highlighted by the NICE Guidelines for Antenatal and Postnatal Mental Health (2007). This is represented in the current project by the relatively low severity of problems identified and the allocation of clients to care clusters not usually associated with the Psychological Medicine CAG within which it is placed.

6. REFERENCES

- Amin, S., Singh, S. P., Croudace, T., et al. (1999). Evaluating the Health of the Nation Outcome Scales. Reliability and validity in a three-year follow-up of first-onset psychosis. *British Journal of Psychiatry*, 174, 399-403.
- Audin, K., Margison, F., Mellor-Clark, J., & Barkham, M. (2001). Value of HoNOS in assessing patient change in NHS psychotherapy and psychological treatment services. *British Journal of Psychiatry*, 178, 561-566.
- Austin, M-P. (2004). Antenatal screening and early intervention for “perinatal” distress, depression and anxiety: where to from here? *Archives of Women's Mental Health*, 7 (1), 1-6.
- Austin, M-P. (2010). Classification of mental health disorders in the perinatal period: future directions for DSM-V and ICD-11. *Archives of Women's Mental Health*, 13 (1), 41-44.
- CEMACH, 2009. Confidential Enquiry into Maternal and Child Health (CEMACH). *Perinatal mortality 2007: United Kingdom*. London.
- Clinical Outcome for Routine Evaluation – Outcome Measure (CORE-OM). [online]. Available at <<http://www.coreims-online.co.uk>> [Accessed 30 July 2010]
- Department of Health (2009). *New Horizons consultation document*.
- Department of Health. (2008). *High Quality Care For All: NHS Next Stage Review Final Report*. CM 7432. London, Stationary Office.
- Green, L., & Hunt, D. (2006). *SLaM Policy for the Care and Support of Pregnant Women with a Diagnosis of Severe Mental Illness*. South London & Maudsley NHS Trust.
- Hay, D.F., Pawlby, S., Angold, A., Harold, G.T., & Sharp, D. (2003). Pathways to violence in the children of depressed mothers. *Developmental Psychology*, 39, 1083-1094.
- Hay, D.F., Pawlby, S., Sharp, D., Asten, P., Mills, A., & Kumar, R. (2001). Intellectual problems shown by 11-year old children whose mothers had postnatal depression. *Journal of Child Psychology and Psychiatry*, 42, 871-889.
- Jacobson, N.S., & Truax, P. (1991). Clinical Significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.

Jayawant, S., Rawlinson, A., Gibbon, F., Price, J., Schulte, J., Sharples, P., Sibert, J.R., & Kemp, A.M. (1998). Subdural haemorrhages in infants: population based study. *British Medical Journal*, 317, 1558–61.

Kendal, R.E., Chalmers, J.C., & Platz, C. (1987). Epidemiology and puerperal psychoses. *British Journal of Psychiatry* 150, 662–673.

Kim-Cohen, J., Moffitt, T.E., Taylor, A., Pawlby, S.J., & Caspi, A. (2005). Maternal Depression and Children's Antisocial Behavior: Nature and Nurture Effects. *Archives of General Psychiatry*, 62 (2), 173-181.

Marks, M., & Kumar, R. (1993). Infanticide in England and Wales. *Medicine Science and the Law*, 33 (4), 329-339.

Marks, M.N., Wieck, A., Checkley, S.A., et al. (1992). Contribution of psychological and social factors to psychotic and non-psychotic relapse after childbirth in women with previous histories of affective disorder. *Journal of Affective Disorders*, 29, 253-264.

Mental Health Outcomes Compendium, National Institute for Mental Health in England, 2008.

Mycroft, R. (2009) Audit of Croydon Perinatal CMHT Caseload 23/04/2009.

National Institute of Clinical Excellence (2007). *Antenatal and Postnatal Mental Health: The NICE Guideline on Clinical Management and Service Guidance*.

Office for National Statistics (ONS). (2009) *Birth summary tables, England and Wales 2009 - Table 4: Live births by local authority of usual residence of mother, numbers, General Fertility Rates and Total Fertility Rates, 2009*. [online]. Available at: <<http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14408>> [Accessed 30 July 2010]

Orrell, M., Yard, P., Handysides, J., & Schapira, R. (1999). Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. *British Journal of Psychiatry*, 174, 409-412.

Parabiaghi, A., Barbato, A., D'avanzo, B., Erlicher, A & Lora, A. (2005). Assessing reliable and clinically significant change on Health of the Nation Outcome Scales: method for displaying longitudinal data. *Australian and New Zealand Journal of Psychiatry*, 39 (8), 719-724.

Rees, A., Richards, A., & Sharpiro, D. (2004). Utility of the HoNOS in measuring change in a Community Mental Health Care population. *Journal of Mental Health*, 13 (3), 295-304.

Self, R., Rigby, A., Leggett, C., & Paxton, R. (2008). Clinical Decision Support Tool: A rational needs-based approach to making clinical decisions. *Journal of Mental Health*, 17(1), 33-48.

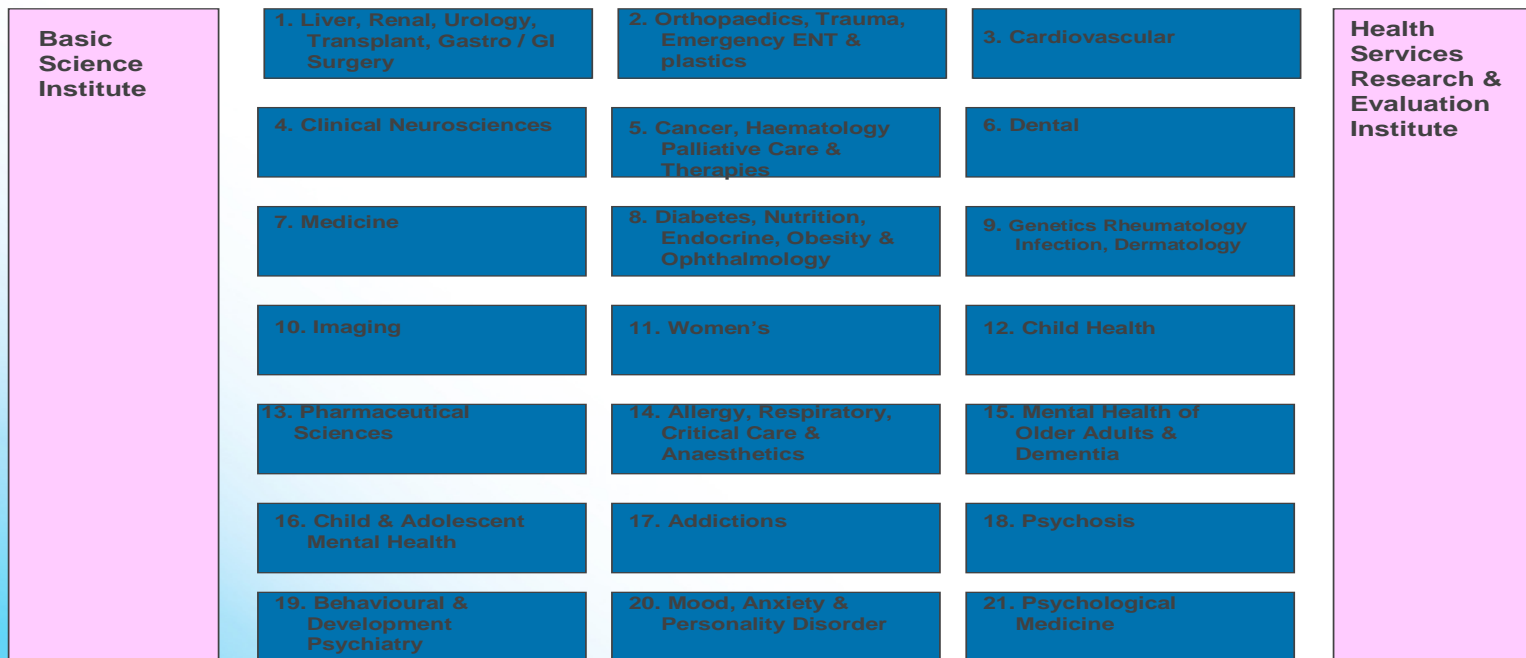
Slade, M., Thornicroft, A., Beck, A., Blindman, J., & Wright, S. (2000). HoNOS Update. Assessment: From Theory to Practice. *The British Journal of Psychiatry*, 176, 394-395.

Wing, J.K., Curtis, R.H., & Beevor, A. S. (1999). Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry*, 174 (5), 432-43.

7. APPENDICES

APPENDIX 1

King's Health Partners Clinical Academic Group Structure (September 2009)



(NB Pathology remains to be resolved)

APPENDIX 2

Mental Health Clustering Tool (MHCT) version 1.0 (January 2010)

The MHCT incorporates items from the Health of the Nations Outcome Scales (HoNOS) (Wing et al. 1999²) and the Summary of Assessments of Risk and Need (SARN) (Self et al 2008¹) in order to provide all the information necessary to allocate individuals to clusters.

HoNOS is an internationally recognised outcome measure developed by the Royal College of Psychiatrists Research Unit (CRU) to measure health and social functioning outcomes in mental health services. The aim of the HoNOS was to produce a brief measure capable of being completed routinely by clinicians and recorded as part of a minimum mental health dataset. The first twelve items of the MHCT are HoNOS items. The HoNOS items are used here with the permission of the Royal College of Psychiatrists, who hold the copyright.

SARN

The Summary of Assessments of Risk and Need (SARN) has been developed by the Care Pathways and Packages Project to aid in the process of establishing a classification of service users based on their needs so that appropriate service responses could be developed both at the individual and service level. It provides a brief description of the needs of people entering into Mental Health Services for the first time or presenting with a possible need for change in their care or treatment. It allows professionals from a range of backgrounds to summarise their assessments in a shared format. Thus it provides a common language for describing health states and related social conditions and improves communication between different users of the tool including health and social care professionals, service users themselves, commissioners and researchers.

Mental Health Clustering Tool (MHCT)

Part 1 contains scales relating to the severity of problems experienced by the individual during the 2 weeks prior to the date of the rating.

Part 2 contains scales that consider problems from a 'historical' perspective. These will be problems that occur in episodic or unpredictable ways. Whilst they may not have been experienced by the individual during the two weeks prior to the rating date, clinical judgement would suggest that there is still a cause for concern that cannot be disregarded (i.e. no evidence to suggest that the person has changed since the last occurrence either as a result of time, therapy, medication or environment etc.) In these circumstances, any event that remains relevant to the cluster allocation (and hence the interventions offered) should be included.

Summary of rating information

- 1) Rate each scale in order from 1 to 13, followed by A to E in part 2.
 - 2) Do not include information rated in an earlier scale except for scale 10 which is an overall rating.
 - 3) Rate the **MOST SEVERE** problem that occurred
 - 4) All scales follow the format:
 - 0 = no problem
 - 1 = minor problem requiring no action
 - 2 = mild problem but definitely present
 - 3 = moderately severe problem
 - 4 = severe to very severe problem
- Rate 9 if Not Known but be aware that this is likely to make accurate clustering impractical.*

N.B. The first data item (current rating of Overactive, aggressive, disruptive or agitated behaviour) is not used in the clustering process, hence does not appear on the cluster profiles. All other ratings are used.

References

¹Self R; Rigby A; Leggett C and Paxton R (2008) Clinical Decision Support Tool: A rational needs-based approach to making clinical decisions. *Journal of Mental Health*, 17(1): 33-48.

²Wing, J. K., Curtis, R. H. & Beevor, A. S. (1999) Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry*, 174 (5), 432-434.

PART 1:

Current Ratings

For scales 1-13, rate the most severe occurrence in the previous two weeks

1. Overactive, aggressive, disruptive or agitated behaviour (current)	
<ul style="list-style-type: none"> • Include such behaviour due to any cause (eg drugs, alcohol, dementia, psychosis, depression, etc). • Do not include bizarre behaviour rated at Scale 6. 	<p>0 No problem of this kind during the period rated.</p> <p>1 Irritability, quarrels, restlessness etc not requiring action.</p> <p>2 Includes aggressive gestures, pushing or pestering others; threats or verbal aggression; lesser damage to property (eg broken cup, window); marked over-activity or agitation.</p> <p>3 Physically aggressive to others or animals (short of rating 4); threatening manner; more serious over-activity or destruction of property.</p> <p>4 At least one serious physical attack on others or on animals; destructive of property (eg fire-setting); serious intimidation or obscene behaviour. <i>Rate 9 if not known</i></p>
2. Non-accidental self-injury (current)	
<ul style="list-style-type: none"> • Do not include accidental self-injury (due eg to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and the injury at Scale 5. • Do not include illness or injury as a direct consequence of drug/alcohol use rated at Scale 3 (eg cirrhosis of the liver or injury resulting from drink driving are rated at Scale 5). 	<p>0 No problem of this kind during the period rated.</p> <p>1 Fleeting thoughts about ending it all but little risk during the period rated; no self-harm.</p> <p>2 Mild risk during the period rated; includes non-hazardous self-harm (eg wrist-scratching).</p> <p>3 Moderate to serious risk of deliberate self-harm during the period rated; includes preparatory acts (eg collecting tablets).</p> <p>4 Serious suicidal attempt and/or serious deliberate self-injury during the period rated. <i>Rate 9 if Not Known</i></p>
3. Problem-drinking or drug-taking (current)	
<ul style="list-style-type: none"> • Do not include aggressive/destructive behaviour due to alcohol or drug use, rated at Scale 1. • Do not include physical illness or disability due to alcohol or drug use, rated at Scale 5. 	<p>0 No problem of this kind during the period rated.</p> <p>1 Some over-indulgence but within social norm.</p> <p>2 Loss of control of drinking or drug-taking, but not seriously addicted.</p> <p>3 Marked craving or dependence on alcohol or drugs with frequent loss of control; risk taking under the influence.</p> <p>4 Incapacitated by alcohol/drug problem. <i>Rate 9 if Not Known</i></p>
4. Cognitive problems (current)	
<ul style="list-style-type: none"> • Include problems of memory, orientation and understanding associated with any disorder: learning disability, dementia, schizophrenia, etc • Do not include temporary problems (eg hangovers) resulting from drug/alcohol use, rated at Scale 3. 	<p>0 No problem of this kind during the period rated.</p> <p>1 Minor problems with memory or understanding (eg forgets names occasionally).</p> <p>2 Mild but definite problems (eg has lost the way in a familiar place or failed to recognise a familiar person); sometimes mixed up about simple decisions.</p> <p>3 Marked disorientation in time, place or person; bewildered by everyday events; speech is sometimes incoherent; mental slowing.</p> <p>4 Severe disorientation (eg unable to recognise relatives); at risk of accidents; speech incomprehensible; clouding or stupor. <i>Rate 9 if Not Known</i></p>

<p>5. Physical illness or disability problems (current)</p> <ul style="list-style-type: none"> • Include illness or disability from any cause that limits or prevents movement, or impairs sight or hearing, or otherwise interferes with personal functioning. • Include side-effects from medication; effects of drug/alcohol use; physical disabilities resulting from accidents or self-harm associated with cognitive problems, drink-driving, etc • Do not include mental/behavioural problems rated at Scale 4. 	<p>0 No physical health problem during the period rated.</p> <p>1 Minor health problems during the period (eg cold, non-serious fall, etc).</p> <p>2 Physical health problem imposes mild restriction on mobility and activity.</p> <p>3 Moderate degree of restriction on activity due to physical health problem.</p> <p>4 Severe or complete incapacity due to physical health problem. <i>Rate 9 if Not Known</i></p>
<p>6. Problems associated with hallucinations and delusions (current)</p> <ul style="list-style-type: none"> • Include hallucinations and delusions irrespective of diagnosis. • Include odd and bizarre behaviour associated with hallucinations or delusions. • Do not include aggressive, destructive or overactive behaviours attributed to hallucinations or delusions, rated at Scale 1. 	<p>0 No evidence of hallucinations or delusions during the period rated.</p> <p>1 Somewhat odd or eccentric beliefs not in keeping with cultural norms.</p> <p>2 Delusions or hallucinations (eg voices, visions) are present, but there is little distress to patient or manifestation in bizarre behaviour, ie clinically present but mild.</p> <p>3 Marked preoccupation with delusions or hallucinations, causing much distress and/or manifested in obviously bizarre behaviour, ie moderately severe clinical problem.</p> <p>4 Mental state and behaviour is seriously and adversely affected by delusions or hallucinations, with severe impact on patient. <i>Rate 9 if Not Known</i></p>
<p>7. Problems with depressed mood (current)</p> <ul style="list-style-type: none"> • Do not include over-activity or agitation, rated at Scale 1. • Do not include suicidal ideation or attempts, rated at Scale 2. • Do not include delusions or hallucinations, rated at Scale 6. 	<p>0 No problem associated with depressed mood during the period rated.</p> <p>1 Gloomy; or minor changes in mood.</p> <p>2 Mild but definite depression and distress (eg feelings of guilt; loss of self-esteem).</p> <p>3 Depression with inappropriate self-blame; preoccupied with feelings of guilt.</p> <p>4 Severe or very severe depression, with guilt or self-accusation. <i>Rate 9 if Not Known</i></p>
<p>8. Other mental and behavioural problems (current)</p> <ul style="list-style-type: none"> • Rate only the most severe clinical problem not considered at scales 6 and 7 as follows. • Specify the type of problem by entering the appropriate letter: A phobic; B anxiety; C obsessive-compulsive; D mental strain/tension; E dissociative; F somatoform; G eating; H sleep; I sexual; J other, specify. 	<p>0 No evidence of any of these problems during period rated.</p> <p>1 Minor problems only.</p> <p>2 A problem is clinically present at a mild level (eg patient has a degree of control).</p> <p>3 Occasional severe attack or distress, with loss of control (eg has to avoid anxiety provoking situations altogether, call in a neighbour to help, etc) ie moderately severe level of problem.</p> <p>4 Severe problem dominates most activities. <i>Rate 9 if Not Known</i></p>
<p>9. Problems with relationships (current)</p> <ul style="list-style-type: none"> • Rate the patient's most severe problem associated with active or passive withdrawal from social relationships, and/or non-supportive, destructive or self-damaging relationships. 	<p>0 No significant problem during the period.</p> <p>1 Minor non-clinical problems.</p> <p>2 Definite problem in making or sustaining supportive relationships; patient complains and/or problems are evident to others.</p> <p>3 Persisting major problem due to active or passive withdrawal from social relationships and/or to relationships that provide little or no comfort or support.</p> <p>4 Severe and distressing social isolation due to inability to communicate socially and/or withdrawal from social relationships. <i>Rate 9 if Not Known</i></p>

10. Problems with activities of daily living (current)	
<ul style="list-style-type: none"> Rate the overall level of functioning in activities of daily living (ADL) (eg problems with basic activities of self-care such as eating, washing, dressing, toilet; also complex skills such as budgeting, organising where to live, occupation and recreation, mobility and use of transport, shopping, self-development, etc). Include any lack of motivation for using self-help opportunities, since this contributes to a lower overall level of functioning. Do not include lack of opportunities for exercising intact abilities and skills, rated at Scales 11-12. 	<p>0 No problem during period rated; good ability to function in all areas.</p> <p>1 Minor problems only (eg untidy, disorganised).</p> <p>2 Self-care adequate, but major lack of performance of one or more complex skills (see above).</p> <p>3 Major problem in one or more areas of self-care (eating, washing, dressing, toilet) as well as major inability to perform several complex skills.</p> <p>4 Severe disability or incapacity in all or nearly all areas of self-care and complex skills. <i>Rate 9 if Not Known</i></p>
11. Problems with living conditions (current)	
<ul style="list-style-type: none"> Rate the overall severity of problems with the quality of living conditions and daily domestic routine. Are the basic necessities met (heat, light, hygiene)? If so, is there help to cope with disabilities and a choice of opportunities to use skills and develop new ones? Do not rate the level of functional disability itself, rated at Scale 10. <p>NB: Rate patient's usual situation. If in acute ward, rate activities during period before admission. If information not available, rate 9.</p>	<p>0 Accommodation and living conditions are acceptable; helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help.</p> <p>1 Accommodation is reasonably acceptable although there are minor or transient problems (eg not ideal location, not preferred option, doesn't like the food, etc).</p> <p>2 Significant problem with one or more aspects of the accommodation and/or regime (eg restricted choice; staff or household have little understanding of how to limit disability or how to help use or develop new or intact skills).</p> <p>3 Distressing multiple problems with accommodation (eg some basic necessities absent); housing environment has minimal or no facilities to improve patient's independence.</p> <p>4 Accommodation is unacceptable (eg lack of basic necessities, patient is at risk of eviction, or 'roofless', or living conditions are otherwise intolerable) making patient's problems worse. <i>Rate 9 if Not Known</i></p>
12. Problems with occupation and activities (current)	
<ul style="list-style-type: none"> Rate the overall level of problems with quality of day-time environment. Is there help to cope with disabilities, and opportunities for maintaining or improving occupational and recreational skills and activities? Consider factors such as stigma, lack of qualified staff, access to supportive facilities eg staffing and equipment of day centres, workshops, social clubs, etc. Do not rate the level of functional disability itself, rated at Scale 10. <p>NB: Rate patient's usual situation. If in acute ward, rate activities during period before admission. If information not available, rate 9.</p>	<p>0 Patient's day-time environment is acceptable: helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self help.</p> <p>1 Minor or temporary problems (eg late giro cheques); reasonable facilities available but not always at desired times, etc</p> <p>2 Limited choice of activities; lack of reasonable tolerance (eg unfairly refused entry to public library or baths, etc); handicapped by lack of a permanent address; insufficient carer or professional support; helpful day setting available but for very limited hours.</p> <p>3 Marked deficiency in skilled services available to help minimise level of existing disability; no opportunities to use intact skills or add new ones; unskilled care difficult to access.</p> <p>4 Lack of any opportunity for daytime activities makes patient's problems worse. <i>Rate 9 if Not Known</i></p>
13. Strong unreasonable beliefs occurring in non-psychotic disorders only. (current)	
<ul style="list-style-type: none"> Rate any apparent strong unreasonable beliefs (found in some people with disorders such as Obsessive Compulsive Disorder, Anorexia Nervosa, personality disorder, morbid jealousy etc) Do not include Delusions rated at scale 6. Do not include Severity of disorders listed above where strong unreasonable beliefs are not present – rated at Scale 9. Do not include Beliefs / behaviours consistent with a person's culture. 	<p>0 No Strong unreasonable beliefs evident.</p> <p>1 Holds illogical or unreasonable belief(s) but has insight into their lack of logic or reasonableness and can challenge them most of the time and they have only a minor impact on the individual's life.</p> <p>2 Holds illogical or unreasonable belief(s) but individual has insight into their lack of logic or reasonableness. Belief(s) can be successfully challenged by individual on occasions. Does not have a significant negative impact on the person's life.</p> <p>3 Holds strong illogical and unreasonable belief(s) but has some insight into the relationship between the beliefs and the disorder. Belief(s) can be 'shaken' by rational argument. Tries to resist belief but with little effect. Has a significant negative impact on person's life. The disorder makes treatment more difficult than usual.</p> <p>4 Holds strong illogical or unreasonable belief(s) with little or no insight in the relationship between the belief and the disorder. Belief(s) cannot be 'shaken' by rational argument. Does not attempt to resist belief(s). Has a significant negative impact on the person's life or other people's lives and the disorder is very resistant to treatment. <i>Rate 9 if not known</i></p>

PART 2:

Historical Ratings

For scales A-E, rate problems that occur in an episodic or unpredictable way. Whilst there may not be any direct observation or report of a manifestation during the last two weeks the evidence and clinical judgement would suggest that there is still a cause for concern that cannot be disregarded (ie no evidence to suggest that the person has changed since the last occurrence either as a result of time, therapy, medication or environment etc). In these circumstances, any event that remains relevant to the current plan of care should be included.

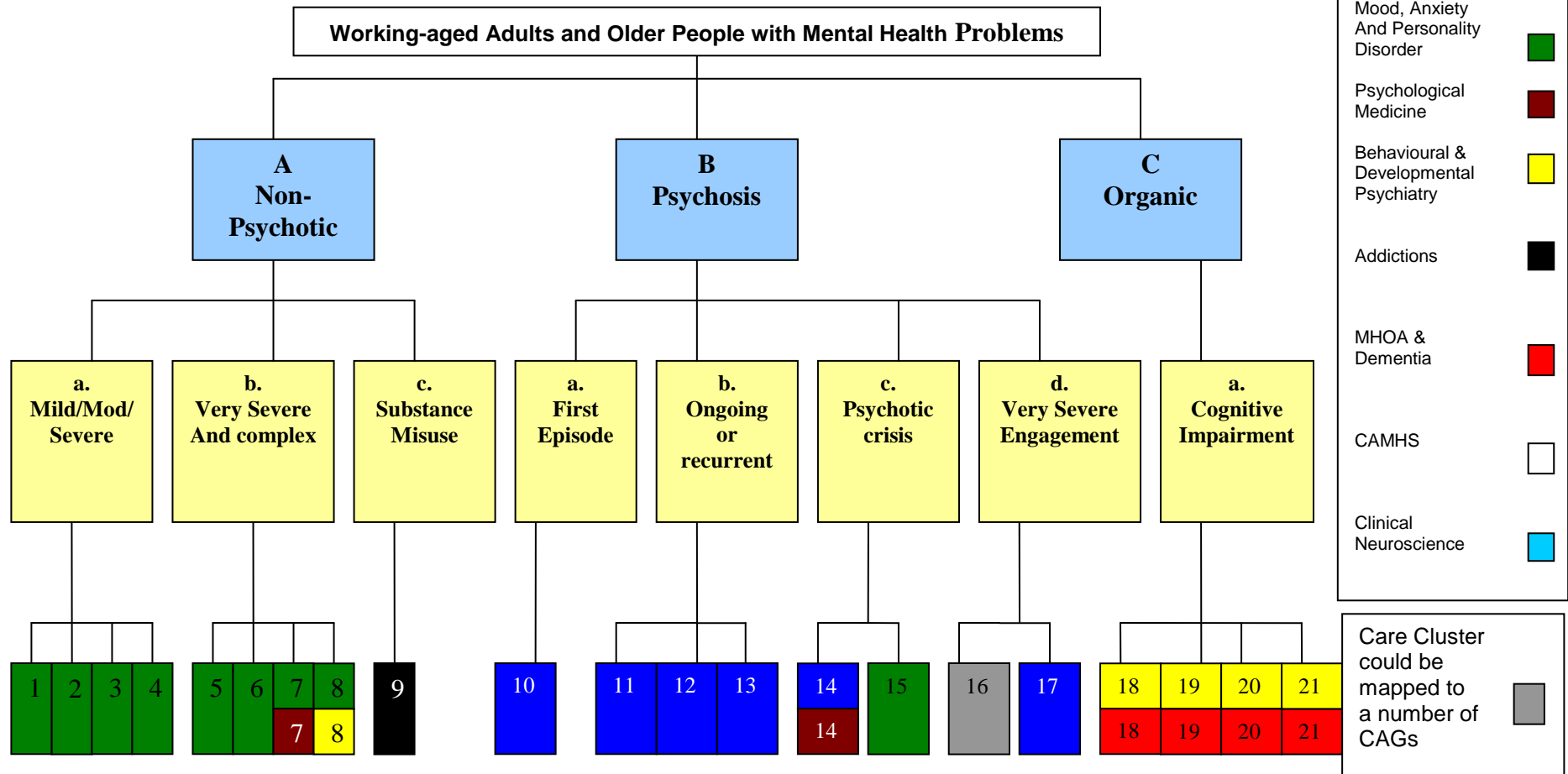
<p>A. Agitated behaviour/ expansive mood (historical)</p> <ul style="list-style-type: none"> Rate agitation and overactive behaviour causing disruption to social role functioning. Behaviour causing concern or harm to others. Elevated mood that is out of proportion to circumstances. Include such behaviour due to any cause (eg drugs, alcohol, dementia, psychosis, depression etc). Excessive irritability, restlessness, intimidation, obscene behaviour and aggression to people animals or property. Do not include odd or bizarre behaviour to be rated at Scale 6. 	<p>0 No needs in this area.</p> <p>1 Presents as irritable, argumentative with some agitation. Some signs of elevated mood or agitation not causing disruption to functioning.</p> <p>2 Makes verbal/gestural threats. Pushes/pesters but no evidence of intent to cause serious harm. Causes minor damage to property (eg glass or crockery). Is obviously over-active or agitated.</p> <p>3 Agitation or threatening manner causing fear in others. Physical aggression to people or animals. Property destruction. Serious levels of elevated mood, agitation, restlessness causing significant disruption to functioning.</p> <p>4 Serious physical harm caused to persons/animals. Major destruction of property. Seriously intimidating others or exhibiting highly obscene behaviour. Elevated mood, agitation, restlessness causing complete disruption. <i>Rate 9 if not known</i></p>
<p>B. Repeat self-harm (historical)</p> <ul style="list-style-type: none"> Rate repeat acts of self harm with the intention of managing people, stressful situations, emotions or to produce mutilation for any reason. Include self cutting, biting, striking, burning, breaking bones or taking poisonous substances etc. Do not include accidental self-injury (due eg to learning disability or cognitive impairment); the cognitive problem is rated at Scale 4 and the injury at Scale 5. Do not include harm as a direct consequence of drug/alcohol use (eg liver damage) to be rated at Scale 3. Injury sustained whilst intoxicated to be rated at Scale 5. Do not include harm with intention of killing self rated at Scale 2. 	<p>0 No problem of this kind.</p> <p>1 Superficial scratching or non-hazardous doses of drugs.</p> <p>2 Superficial cutting, biting, bruising etc or small ingestions of hazardous substances unlikely to lead to significant harm even if hospital treatment not sought.</p> <p>3 Repeat self-injury requiring hospital treatment. Possible dangers if hospital treatment not sought. However, unlikely to leave lasting severe damage even if behaviour continues providing hospital treatment sought.</p> <p>4 Repeat serious self-injury requiring hospital treatment and likely to leave lasting severe damage if behaviour continues (ie, severe scarring, crippling or damage to internal organ) and possibly to death. <i>Rate 9 if not known</i></p>

C. Safeguarding Children & Vulnerable Dependent Adults (historical)	
<ul style="list-style-type: none"> Rate the potential or actual impact of the patient's mental illness, or behaviour, on the safety and well being of vulnerable and dependent persons, including children, vulnerable adults and dependent elders. Include any patient who lives in a household with children under the age of 18 years. Include any patient who has substantial access and contact with children or other vulnerable persons. Do not include risk to wider population covered at scale 14. Do not include challenge to relations covered in scale 9. 	<p>0 No obvious impact of the individuals' illness or behaviour on the safety or well being of vulnerable persons.</p> <p>1 Mild concerns about the impact of the individual's illness or behaviour on the safety or well-being of vulnerable persons.</p> <p>2 Illness or behaviour has an impact on the safety or well being of vulnerable persons. The individual is aware of the potential impact but is supported and is able to make adequate arrangements.</p> <p>3 Illness or behaviour has an impact on the safety or well being of vulnerable persons but does not meet the criteria to score 4. There may be delusions, suicide risk or self-harm. However, the individual has insight, can take action to significantly reduce the impact of their behaviour on the children and is adequately supported.</p> <p>4 Without action the illness or behaviour is likely to have direct or indirect significant impact on the safety or well-being of vulnerable persons. Problems such as delusions, severe suicide risk or problems of impulse control may be present. There may be lack of insight, an inability or unwillingness to take precautions to protect vulnerable persons and/or lack of adequate support and protection for vulnerable persons. <i>Rate 9 if not known</i></p>
D. Engagement (historical)	
<ul style="list-style-type: none"> Rate the individual's motivation and understanding of their problems, acceptance of their care/treatment and ability to relate to care staff. Include the ability, willingness or motivation to engage in their care/ treatment appropriately, agreeing personal goals, attending appointments. Dependency issues. Do not include Cognitive issues as in scale 4, severity of illness or failure to comply due to practical reasons. 	<p>0 Has ability to engage/disengage appropriately with services. Has good understanding of problems and care plan.</p> <p>1 Some reluctance to engage or slight risk of dependency. Has understanding of own problems.</p> <p>2 Occasional difficulties in engagement ie missed appointments or contacting services between appointments inappropriately. Some understanding of own problems.</p> <p>3 Contacts services inappropriately. Has little understanding of own problems. Unreliable attendance at appointments. Or attendance depends on prompting or support.</p> <p>4 Contacts multiple agencies ie GP, A & E etc, constantly. Little or no understanding of own problems. Fails to comply with planned care. Rarely attends appointments. Refuses service input. Or Attendance and compliance dependent on intensive prompting and support. <i>Rate 9 if not known</i></p>
E. Vulnerability (historical)	
<ul style="list-style-type: none"> Rate failure of an individual to protect themselves from risk of harm to their health and safety or well-being. Include physical, sexual, emotional and financial exploitation or harm/ harassment Do not include problems of engagement rated at scale D. 	<p>0 No vulnerability evident.</p> <p>1 No significant impact on person's health, safety or well-being.</p> <p>2 Concern about the individual's ability to protect their health, safety or well-being requiring support or removal of existing support would increase concern.</p> <p>3 Clear evidence of significant vulnerability affecting the individuals ability to protect their health and safety or well-being that requires support (but not as severe as a score of 4). Or removal of existing support would increase risk.</p> <p>4 Severe vulnerability – total breakdown in individual's ability to protect themselves resulting in major risk to the individual's health, safety or well-being. <i>Rate 9 if not known</i></p>

Item	Score	Item	Score
Part 1 (Current)		Part 2 (Historical)	
1		A	
2		B	
3		C	
4		D	
5		E	
6			
7			
8			
Please Circle	A B C D E F G H I J		
N.B. If J – (other) please specify			
9			
10			
11			
12			
13			

APPENDIX 3

DECISION TREE [RELATIONSHIP OF CARE CLUSTERS TO EACH OTHER]



APPENDIX 4

Distribution of Care Cluster Identified for Clients Per Caseholder

